76-28,889

MOSS, Abigail Miriam Wender, 1928-
THE DEVELOPMENTAL EFFECTS OF IRRELEVANT
INFORMATION ON PROBLEM SOLVING STRATEGIES.

City University of New York, Ph.D., 1976
Psychology, experimental

Xerox University Microfilms, Ann Arbor, Michigan 48106

The Developmental Effects of Irrelevant Information on Problem Solving Strategies

by

Abigail Moss

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

July 23, 1976
date

Barry Gholson
Chairman of Examining Committee

August 15, 1976
date

Florence L. Denmark
Executive Officer

Supervisory Committee

The City University of New York

ABSTRACT

The Developmental Effects of Irrelevant Information on

Problem Solving Strategies

by

Abigail Moss

Advisor: Professor Barry Gholson

This study investigates the developmental aspects of selective attention to relevant information in the context of a complex problem solving task. In a number of studies, Hagan and his associates (1967, 1968, 1969, 1970) have investigated this issue in relation to serial memory. They found that the recall of central (relevant) information increased with age. Although recall of incidental (irrelevant) information was above chance for all ages, the proportion of incidental to central information recalled, remained constant or dropped between the ages of 12 and 14.

The ability to exclude irrelevant information is critical to complex problem solution. The present study examines the ability of third and eighth graders to process information relevant to problem solution and to employ strategic approaches to solution. Studying children assumed to differ in cognitive capacities should yield a better understanding of the processes involved in the development of selective attention.

Ninety-six children from each grade were individually presented with five practice and eight experimental problems based on the hypothesis testing paradigm (Gholson et al, 1972) with the modification that the blank trials were replaced by verbal probes (Karpf & Levine, 1971). Children in each grade were equally divided into four experimental conditions (12 boys and 12 girls per condition). The conditions differed

in the number of dimensions (four or eight) and whether or not a cue card was provided as a memory aid.

The four conditions were:

1) Four-Dimension-No-Cue - Standard four dimensional bivalued configuration (Gholson et al., 1972). The child was told that all dimensions in the configuration were relevant to solution.

2) Four-Dimension-Cue - -- Same as 1 with the addition of a cue card listing the eight hypotheses that were potential solutions (e.g., black, white, circle, square, big, little, S, P).

3) Eight-Dimension-No-Cue - Eight dimensional bivalued configurations constructed in the same manner as Four-Dimensional-No-Cue. For each problem, the child was told which <u>four</u> of the dimensions were <u>relevant</u> to solution. (In order to solve, the child must relate feedback information to <u>only</u> the designated four.)

4) Eight-Dimension-Cue - Same as 3 with the addition of a cue card listing the eight hypotheses that were potential solutions to that problem.

Age had a significant effect on all aspects of problem solving. Across all conditions, eighth graders consistently solved a larger proportion of the experimental problems and more efficiently processed feedback information (e.g. retained hypotheses following positive feedback, etc.). In the Eight-Dimension-No-Cue condition, both age groups tested irrelevant hypotheses but the younger group did this on significantly more problems (59% vs. 45%). The memory aid (Eight-Dimension-Cue) significantly reduced the proportion of problems showing irrelevant hypotheses in both age groups but the younger was less successful than the older in eliminating them from processing (31% of the problems vs 12%).

The data were analyzed for the effects of age and experimental condition on the problem solving strategies identified by Gholson et al. (1972). Under standard Four-Dimension-No-Cue conditions, the strategies of third graders followed a similar pattern to that of the Gholson et al. (1972) elementary school groups in that the predominant strategic approach was dimension checking (sequential elimination of one dimension per negative feedback trial). However, with the use of a verbal probe, there was a slightly greater proportion of problems in which the most efficient strategy, fousing, was used (simultaneous processing of feedback information to the total set of relevant hypotheses on each feedback trial). The strategy patterm of eighth graders was similar to that of the adults in the Gholson et al. (1972) study. They used an approximately equal proportion of focusing and dimension checking in their problems. Providing a memory aid (Four-Dimension-Cue) did not affect the performance of eighth graders. In contrast, the third graders processed information more efficiently and showed a larger proportion of dimension checking patterns than the Four-Dimension-No-Cue group. The proportion of focusing was not affected. In the Eight-Dimension-No-Cue condition (in comparison with Four-Dimension-No-Cue) problem solving efficiency was decreased in both age groups. Both used a focusing strategy less often and a dimension checking strategy on a greater proportion of their problems. While the addition of a memory aid (Eight-Dimension-Cue) decreased the proportion of problems on which irrelevant hypotheses were included in the sampling set, it did not affect problem solving efficiency in either age group. The third graders focused on a smaller proportion of their problems and dimension checked on the same proportion of problems as the Eight-Dimension-No-Cue group. Eighth graders, on the other hand, dimension checked on more problems and focused on a concomitant smaller number of problems than their peers in the Eight-Dimension-No-Cue condition.

The verbal probe and the cue card are interpreted as reducing the memory load in different ways, thus resulting in increased problem solving efficiency in third-grade children under Four-Dimension-Cue and No-Cue conditions. The introduction of irrelevant information generally decreased efficiency, probably because of the increased demands placed upon attention and memory. In addition, the cue card had differential effects on performance depending upon age and the presence or absence of irrelevant information. The data from the study suggest that problem solving efficiency in this paradigm must be examined in terms of the attentional and memorial demands of the particular stimulus situation as well as of the particular strategy.

I should like to express my appreciation to the members of my committee: Dr. Barry Gholson, Dr. Harry Beilin and Dr. Joseph Glick. In particular, I would like to thank my chairman, Dr. Gholson for his unflagging interest, assistance and support from inception to completion of this project. I should also like to thank Dr. Solomon Weinstock, who generously gave his expertise and time to the statistical analyses and Dr. Marion Gindes, whose interest and critical comments were an invaluable aid at every stage of this research.

In addition, I wish to thank the administration, faculty and students of the Hewlett-Woodmere School District of New York, especially Shelly Kislik, Rolland Beardsley, Alfred Walling, Gerald Glasser and Mildred David for their cooperation, which made this study possible.

I should like to extend special thanks to my husband, Paul and sons, Adam and Daniel for their support, encouragement and understanding throughout the preparation of this dissertation.

Table of Contents

Tables

Figures

Page

The broad area of interest in this study is the development of conceptual thinking in middle childhood and adolescence. The focus is on strategies used in solving concept identification problems. Specifically, the experimental question concerns the effect of irrelevant information on the problem solving behavior of children assumed to be functioning at different cognitive levels.

It is generally agreed that there are limitations on the amount of information the human organism can attend to at any one time (Broadbent, 1958; Hagen, 1967; Moray, 1970; Neisser, 1967; Simon, 1972). While theoretical suppositions regarding the mechanisms and processes involved differ, it is universally accepted that the human adult has the ability to efficiently, if not perfectly, attend to some properties of the environment to the relative exclusion of others (See Egeth, 1966; Moray, 1970; Neisser, 1967; for theoretical reviews).

It is difficult to imagine any activity in which this ability is not an asset, if not a necessity. In order to perform any task, the features of the environment that convey useful information must be differentiated from those which, at best, are not informative and at worst, interfere with performance. Attention must be focused on the information that is relevant.

Observations of young children suggest they make less distinction between relevant and irrelevant aspects of the environment than do older children and adults. The ability to screen-out irrelevant information and selectively attend to relevant information may be relatively deficient in younger children.

Hagen and his associates (Druker & Hagen, 1969; Hagen, 1967; Hagen & Frisch, 1968; Hagen, Meacham & Mesibov, 1970; Hagen & Sabo, 1967; Sabo & Hagen, 1973) have conducted a series of studies that comprise a systematic investigation of the development of stimulus selection and exclusion in the context of a retention task. The studies cover the age range from first grade through college.

The general experimental procedure has been to present a series of visual arrays. Each array consisted of a number of cards. On each card, the stimulus complex consisted of two vertically arranged drawings, each an exemplar of one of two conceptual categories (e.g., animals and household objects). Across the series, the cards varied in location in the array, but not in content. The same two exemplars were always paired (e.g., pig and telephone). The instructional set was to remember the location of the exemplars of one category (e.g., the animals). Following each array, the measure of central (relevant) information recall was the identification of the location of a particular exemplar in the previous array (location-matching). Following the series, incidental (irrelevant) recall was assessed by requiring the subject to match each central exemplar with the exemplar from the other category with which it had been paired in the presentations (pair-matching).

The most general finding across studies was that recall of central information increased significantly with age. Although recall of incidental information was significantly above chance for all age groups, it either remained constant or decreased with age, depending upon the particular task demands. In all studies, the proportion of recalled central information to incidental information began to increase somewhere between the sixth and seventh grades. The drop, at about this age, in the proportion of incidental information retained has also been observed in other contexts. In a modified discrimination learning paradigm, Siegal and Stevenson (1966) found that the amount of incidental learning increased between about the seventh and twelfth years and began to decline between ages twelve and fourteen. Similarly, the incidental information recalled from film presentations increased through the elementary school years and then dropped at the seventh-grade (Collins, 1970; Hale, Miller & Stevenson, 1968). In an investigation of selective listening with distraction, Doyle (1973) found that the retention of central information increased from eight to fourteen years while the retention of distracting information decreased.

Although these studies are few in number, they do include a variety of situational contexts. One observation common to all of them is that within the 11 to 14 age range there is an increase in the proportion of central to incidental information recalled. In the Hagen studies, these findings remained stable despite manipulations of stimulus discriminability (Druker & Hagen, 1969; Hagen & Frisch, 1968); the amount of exposure to incidental information (Baker in Hagen, 1972); and variations in the designation of incidental and central information (Hagen & Sabo, 1967).

Apparently, when the older children were presented with a situation in which they were instructed to attend to a particular set of stimuli, they directed their attention to that set and retained less information about other aspects of the stimulus situation. On the other hand, the younger children did not as effectively direct their attention as evidenced by the finding that there was not as large a difference between the amount of relevant and irrelevant information recalled

(Druker & Hagen, 1969; Hagen, 1967; Hagen & Sabo, 1967; Maccoby & Hagen, 1965).

This raises the issue of how the selection process operates. What is responsible for the change with age in the retention of information from one aspect of a situation to the relative exclusion of information from other environmental sources?

One possibility is that observed changes in functioning are primarily a reflection of physiological development. This view is based on Broadbent's (1958) filter model. Broadbent conceptualized the human nervous system as limited in communication channel capacity. Within that system, filters operate to separate relevant from irrelevant information. The irrelevant information is either rejected at input or enters short term memory and fades without further processing. Only relevant information affects response or is stored in long-term memory.

Hagen (1967) based a developmental model on this conceptualization. The assumption was that the lack of selectivity observed in the behavior of young children was a function of an inadequately developed filter system. Since the theory postulates that the filters function to prevent information overload, it was predicted that the proportion of incidental to central information retained would decrease with age. The expectation was that the older subjects, under information overload conditions, would filter out the incidental information in order to maintain performance on the central memory task. In general, the data did not support this hypothesis. Under those conditions, central task performance decreased significantly at all ages (first- through seventh-grade) whereas there was no significant effect on incidental recall (Hagen, 1967; Maccoby & Hagen, 1965).

Hagen (1967) has proposed a developmental translation of Neisser's (1967) theory of focal attention as a better fit to his experimental observations. In the Broadbent (1958) filter model, the crucial factor involves limitations of the system. The theory is quite precise in delineating the mechanisms of the system but it doesn't attempt an explanation of the process of selective attention.

Neisser's (1967) model of 'focal attention' directly addressed itself to that issue. He postulated a two stage theory of attention. In the first stage, labeled 'preattentive', information is primarily handled in a global, undifferentiated fashion. The second stage is characterized by an active analysis of the stimulus situation and a reconstruction which includes only those aspects of the situation selected for further processing. This stage is described as 'analysis by synthesis'.

Hagen (1972) postulated that the improvement in selectivity observed in the adolescent may be a function of 'analysis by synthesis' processing of information. He suggests that this kind of attentional process may not reach full development until preadolescence. He also suggests that the intake of task relevant and irrelevant information may involve separate processes. Irrelevant information intake may not be directly affected by developmental change. When measured, it may decline in comparison to relevant information as a function of interference due to the encoding and storing of task relevant information. It is his contention that the older children's superiority in recognizing and recalling relevant material in situations in which relevant and irrelevant stimulus features were present is because they visually attend to only the relevant features and verbally encode only that information.

In support of this position, Hagen cites pertinent data from several studies. In the Druker and Hagen (1969) and Sabo and Hagen (1973) studies, there were significantly more post-experimental verbal reports of visual scanning and rehearsing of only relevant material from the older children (seventh- and eighth-grade) than from the younger (third-, fourth-, fifth- and sixth grades). In a study by Hagen, Meacham and Mesibov (1970), subjects had to label each central category as the stimuli were presented. Although labeling did not affect central task performance for subjects in grades four, six and eight, it did interfere with the performance of college students. In the same vein, an imposition of a distractor task (counting aloud) significantly lowered the central task scores of seventh-grade children but not third- or fifth-grade children (Sabo & Hagen, 1973). The authors concluded that labeling and counting interfered with spontaneous rehearsal in the older groups in each study but did not affect performance in the younger groups because, presumably, the children in the latter group were not rehearsing.

The evidence he presents from his own studies is not very compelling. There is no way to determine if the verbal reports of visual scanning and verbal encoding activities actually reflected the activities of the subjects. It is equally likely that differences in this measure are indications of age differences in awareness of one's own activities or in specifically articulating those activities (Flavell, 1971). The evidence from the studies using 'counting' and 'labeling' conditions is equally suspect since the designs do not permit the observation of the inferred rehearsal strategies. But other studies (Appel, Cooper, McCarrel, Sims, Knight, Yussen & Flavell, 1972; Belmont & Butterfield, 1971;

Moely, Olson, Holmes & Flavell, 1969; Neimark, Slotnick & Ulrich, 1971) specifically designed to permit observation of the subject's activity in the process of encoding material for recall have demonstrated that children do use mnemonic strategies and these activities become more spontaneous, varied and task appropriate with increasing age.

Hagen's (1971, 1972) argument is that the subject brings to the experimental situation knowledge he has acquired from his experiences in similar situations. If he has mnemonic strategies available that increase the probability of retaining the material and he realizes both that he has them and that they are task appropriate, he will employ them in the situation. The cognitive activity necessary for memorizing the relevant material then precludes attention to irrelevant aspects of the situation and those aspects will, therefore, neither be scanned nor encoded for storage.

Attention, by definition, involves careful observation and concentration. Attention, therefore, in contrast to looking or listening, has an inherently active quality. Neisser (1967) refers to it as 'concentrating the processes of analysis on a selected portion of the field'. It would seem to be always in the service of another activity, be it memorizing, learning or problem solving. In Hagen's studies the task is to memorize one relationship present in the stimulus situation; usually the relationship of particular category representatives to their location in an array. The second relationship, usually of the two pictures that comprise the stimulus complex, is irrelevant to the task as defined by the experimenter. The behavior of the younger child in these studies does not appear to conform completely to Hagen's (1972) developmental version of Neisser's (1967) theory. The evidence suggests younger children also engage in second stage processing. They seem to engage in 'analysis by synthesis' but the activity is not limited to only the designated relevant relationship. They, for example, remember as much and sometimes more of the information involved in both relationships. Furthermore, the verbal reports of the younger children indicate they engage in the same task appropriate activities as the older children (i.e.,visual scanning and rehearsal) but they do not exclude the relationship designated 'irrelevant' from these activities. It is unclear why they do not. It seems unlikely that fifth-grade children do not understand the relatively simple task instructions.

An examination of the data from several of the Hagen studies (Hagen, 1967; Hagen, Meacham & Mesibov, 1970; Hagen & Sabo, 1967) suggests a possible explanation. Perhaps two separate processes are necessary for maximal performance on a selective attention task. The subject may have to actively attend to the relevant information, encode it, and also actively screen out the irrelevant information so that it does not intrude on task relevant encoding activity. Intake and exclusion may be two separate processes. Furthermore, contrary to Hagen's (1972) supposition, the inhibition of an attentional response to designated irrelevant information may also follow a developmental pattern. The data from the Hagen and Sabo (1967) study, for example, suggest that the ability to inhibit responses to aspects of the stimulus situation which are salient but contrary to the task demands may develop with age. This study compared the effect of pair-matching versus location-matching as the central task (relevant information). The instructional set did not affect the usual finding that older children retain a larger proportion of relevant to irrelevant information. This remained invariant regardless of which relationship was designated task relevant.

It appeared, however, that the pair relationship was salient for all age groups, since the proportion of relevant to irrelevant information recalled was higher in each age group when pair-matching was the central task (relevant information) than when location-matching was the central task. The third-grade children did not vary in performance as a function of task demands; that is, they retained relatively the same proportion of information about each relationship whether it was designated 'relevant' or 'irrelevant.' In contrast, the seventh-and ninth-grade children did vary their performance as a function of task demands. Although there was no increase in the number of locations recalled when that relationship was designated 'relevant', there was a decrease in the number of pairs recalled when that relationship was designated 'irrelevant.' It appears that the third-grade children were more affected by the salience of the pair-matching relationship than the older groups. This suggests,then, that the older children employed an exclusion process of some kind which enabled them to discount the effect of the saliency of pair-matching when it was designated task 'irrelevant'.

Further support for the hypothesis that processes involving selective attention to relevant cues and exclusion of irrelevant cues may be relatively independent and even follow different developmental courses is obtained from a comparison of the Hagen (1967) and Hagen, Meacham and Mesibov (1970) studies. Both studies included a comparison of location relationship recall when no irrelevant information was present (one category stimuli) and when irrelevant information was present (standard two category stimuli). In the Hagen (1967) study, the seventh-grade children showed the usual superiority over the younger groups in the proportion of central (relevant) to incidental (irrelevant) information recalled. However, they were similar to the younger groups in recalling more central information when no irrelevant information was present (one category stimuli) than when it was (two category stimuli). In contrast, the college students in the Hagen et al.(1970) study not only retained a greater proportion of relevant to irrelevant information but also recalled the same number of locations regardless of whether irrelevant information was included in the stimulus complex. This suggests that although both the seventh-grade children and the college students used appropriate encoding strategies, the seventh-grade children were still in some undetermined way affected by the presence of irrelevant information in the stimulus complex, whereas the college students were not.

Apparently, focusing attention on the relevant situational components of this task is more difficult than just applying an appropriate encoding strategy to the relevant information. The data from the Hagen and Sabo (1967) study suggest that differences among age groups may be, to some degree, a function of the younger child's failure to focus attention on an externally designated 'relevant' relationship when the 'irrelevant' relationship is more salient. Whether the relatively small difference between the amount of relevant and irrelevant information recalled by the younger children was due to some difficulty in excluding irrelevant information from encoding activities or was due to not engaging in appropriate encoding activities cannot be determined from the data. But the observation that irrelevant information was a distracting attentional factor for seventh-grade children (Hagen, 1967) and yet did not effect the ratio of relevant to irrelevant information recalled suggests that children in this age group engaged in exclusion and encoding activities which enabled them to offset the intrusion of irrelevant information and meet task demands.

A number of researchers in the area of memory development (Corsini, 1971; Flavell, 1971; Hagen, 1971) have maintained that two important factors contribute to the age-related improvement in recall performance. First, the developing child is acquiring a variety of mnemonic strategies. Second, he is also acquiring an understanding of the nature of different problems and a concomitant recognition of the most appropriate application of a particular atrategy to a particular task.

Effective inhibition of an attentional response to designated irrelevant material may develop gradually. Nevertheless, it may not seriously interfere with task performance if the individual can assess the goals of a task; has a repertoire of means to reach those goals; can determine the particular means which will most effectively lead to the goals; and chooses to use those means.

Flavell (1971) has made the point that memory, problem solving and learning are essentially applied cognition. How these processes operate at a particular point in time is fundamentally a function of the cognitive structural organization operative at that stage of development. Perhaps, the change in the ratio of relevant to irrelevant information recalled which is generally observed in the 11 to 13 age range is a function of the development of the ability to accurately assess a task situation in its entirety and plan a means to reach the goal based on an analysis of the inherent possibilities in the situation. That would seem to be a sine qua non for what Flavell (1963) refers to as the most important general property of formal operational thought, the ability to discover the 'real' from entertaining the total set of 'possibles.' This kind of task analysis may not be possible until the child is at least on the threshold of formal operational thinking.

In this situation, if the ability is present, the individual will attempt to reject the irrelevant information and focus on the information that is relevant, since he 'understands' that the irrelevant information is superfluous to the task at hand. If he does not assess the total situation, he may attempt to comply with task demands in a more concrete fashion; i.e., try to remember what was asked for but make no attempt to exclude from attention or encoding those aspects of the stimulus complex which are extraneous to the task.

The younger child's understanding of any task may be quite different from that of the adolescent (Piaget & Inhelder, 1969). Although the younger child might understand task instructions and try to comply with task demands, the activities he engages in will be limited both by the means at his command and his understanding of what is necessary for effective task performance. The means necessary for achieving the goal of a particular task are determined by the nature of the task, but the ability to analyze any task situation in order to apply the appropriate means is general. Competence at the latter level may be a function of the child's presently existing cognitive organization.

The argument is that the way in which an individual deals with irrelevant information present in a stimulus situation is very much dependent on how he analyzes the situation. By the same reasoning, the way he analyzes the situation is basically dependent on his cognitive level.

But the way in which he deals with irrelevant information is also dependent on the nature of the task. In recall studies, the task is to encode a set of data, store them for a short period (or continuously rehearse them) and then retrieve the data and respond appropriately. Depending on information load, attention to the irrelevant information may or may not preclude success on the task; i.e., recalling the relevant data.

A problem solving situation is more complex. It would seem that the elimination of irrelevant information from the data to be processed would be a most important factor in successful problem solving. To efficiently tackle a problem, the initial step would seem to be to consider all the possible relations that might lead to solution. If irrelevant information is included as basic data, the problem would, minimally, involve many more steps to solution and might never be solved. Presenting children of different ages with a problem solving task in which separation and exclusion of irrelevant information are more crucial than in a simple memory task should lead to a better understanding of how children assumed to be functioning on different cognitive levels handle selective attention to relevant information.

A paradigm that seems appropriate is Levine's (1966) hypothesis testing model. Within the last few years, this paradigm has been used in a number of concept identification studies with children (Eimas, 1969, 1970; Gholson, Levine & Phillips, 1972; Gholson, Phillips & Levine, 1973; Ingalls & Dickerson, 1969; Neussle, 1972). The task is usually a four-dimensional bivalued discrimination-learning problem that utilizes a hypothesis probe technique. This technique, commonly involving blank trials, permits the monitoring of a subject's sequence of hypotheses throughout a problem. There are two sets of stimuli; each consists of four pairs of stimulus configurations. In each set, each level of every dimension appears twice with each level of every other dimension. With the use of such counterbalanced stimuli on blank trials, the hypothesis the subject is currently testing can be inferred from the series of response choices on the non-feedback trials. With the use of such counterbalanced stimuli on feedback trials, the problem can be logically solved following the third feedback trial.

The above mentioned studies cover a wide age range (kindergarten through college). Examining the patterns of findings across studies, the following conclusions seem appropriate: From second grade on, under standard conditions, children reliably formulated and used hypotheses. They retained hypotheses when feedback was positive and rejected them when feedback was negative. The findings from these studies lend strong support for the contention that children as young as second grade behave according to the assumptions of Levine's hypothesis theory (see Levine, 1966 for details).

The performances of the children were not, however, identical to that of adults. Gholson et al,(1972) found that the percentage of problems solved increased with age. The largest difference was between sixth-grade children and adults. The efficiency of information processing was measured in the Eimas (1969, 1970), Ingalls and Dickerson (1969), and Neussle (1972) studies by a 'focusing' formula (Levine, 1966). This formula estimates the number of hypotheses in the set from which the subject samples following negative feedback. Eimas (1969) reported that only his college population utilized information efficiently. In the Neussle (1972) study, ninth-grade children processed information more efficiently than fifth-grade children. However, neither age group showed maximum efficiency. Ingalls and Dickerson (1969) reported similar findings. Although processing efficiency increased with age, even college students did not manifest perfect processing

Gholson et al.(1972, 1973) employed a different technique for the evaluation of information processing. This model takes into account the plan the subject follows across a series of feedback trials and hypotheses probes. The technioue for characterizing a plan is analagous to the technique for detecting hypotheses using blank-trial probes. Just as an hypothesis may be inferred from a seouence of responses during a blank-trial probe, the hypothesis sampling system may be inferred from a seouence of hypotheses and feedback trial information. The hypotheses observed following the first three feedback trials are used for system categorization. It is assumed that a system is a determinant of the seouence in which hypotheses are manifested. Based on the data from the original experiments (Gholson et al.,1972), six systems were identified. Three of these were characterized as strategies; these involve seouences of hypotheses dictated by a logical format. They are methods of processing information which, in principle, lead to solution. The other three systems, which involve repeated manifestations of the same hypothesis, have characteristics which do not fit into the information processing conception and thus, were characterized as stereotypes. A subject following such a system would, in principle, never attain solution.

The following is a brief description of the response patterns that define each of these systems and the percentage of problems that Gholson et al.(1972, 1973) observed in each of the categories.

Strategies

Focusing - It is assumed that a subject using a focusing plan begins the problem by simultaneously considering all eight hypotheses as the potential solution. All hypotheses inconsistent with feedback are eliminated following each feedback trial. Each of the first three hypotheses are consistent with all previous feedback information (global consistency, Gregg & Simon, 1967) and solution is reached following the third feedback trial. Adults manifested this strategy in about half of their problems. Ten percent or fewer of the elementary school children's problems could be placed in this category.

Dimension Checking - This is a less efficient strategy. Unlike focusing, it is not assumed that the use of a dimension checking plan entails the retention of all previous feedback information. A subject using a dimension checking plan is assumed to order the hypotheses into pairs on each dimension and to test one dimension at a time, choosing the value of the dimension that is consistent with feedback from only the immediately preceding negative feedback trial. Only one hypothesis from each dimension is manifested since the subject logically infers that its complement was disconfirmed at the time the observed hypothesis was chosen. This strategy was observed in approximately 35% to 45% of the adult problems. Fourth-and sixth-grade children showed this pattern in approximately 55% to 60% of their problems. Under standard conditions, second-grade children showed this strategy in 55% of their problems. There were no developmental trends in the use of this system among the elementary-school groups. Under standard conditions, kindergarten children did not manifest this strategy.

Hypothesis Checking - This is the least efficient strategy. A subject using this plan is also assumed to order hypotheses into pairs on each dimension. But, in contrast to dimension checking, each hypothesis of each dimension is tested in turn: e.g., the subject manifests 'black', then 'white' before rejecting the color dimension, etc. Adults manifested this pattern during fewer than 10% of their problems. Fourth- and sixth-grade children showed this in 5% to 15% of their problems and second-grade children manifested it in about 25%. Kindergarten children, under standard conditions, showed it during about 5% of their problems.

Stereotypes

Stimulus Preference - The first hypothesis sampled is consistently repeated despite disconfirmation by negative feedback. Virtually no adult subject ever manifested this system. Five to fifteen percent of the problems of fourth- and sixth-grade children were placed in this category. Manifestation of this system was very dependent on task conditions for both kindergarten and second-grade children. Within the second-grade groups, 5% to 35% of the problems followed this pattern.

Position Alternation - Positional responses are alternated throughout each blank-trial probe (i.e., left, right, left, right or right, left, right, left). This pattern was primarily observed in the problems of kindergarten children.

Position Perseveration - One positional response is manifested throughout the series of blank-trial probes (i.e. right, right, right, right or left, left, left, left). This system was also observed primarily in kindergarten children.

In addition, a category labeled 'Random' consisted of eight specific response sequences which did not fit into any of the six previously described patterns. These sequences were observed in about 3% of the problems of adults and from 10% to 15% of the problems of elementary-school children.

The techniques described in the Gholson et al. (1972) study permit the observation of the subject's sequence of hypotheses throughout the course of problem solving. The systems analysis appears to provide a useful description of the subject's problem solving behavior. In principle, the problem can be solved by any of three systematic plans (strategies). Each strategy makes a differential demand on the intellectual and mnemonic resources of the subject (see Appendix A for detailed description of the demands of each). Gholson's data suggest that children as young as second-grade can follow a plan in attempting to solve a problem. They systematically use stimulus and feedback information in formulating and testing hypotheses. There are, however, age related differences both in the means (strategies) employed in attempt to reach the goal and successful achievement of the goal. Under standard conditions, kindergarten children nearly always responded to this problem in a stereotypic fashion, ignoring the demands of the task. The three elementary school groups did not differ much in the systems they employed. Primarily, they tested hypotheses, dimension by dimension. But they did differ in successful achievement of the goal (i.e., discovering the solution hypothesis) with success increasing with age. The most efficient strategy, focusing, was rarely observed among children of elemenatry-school age. Adults used this plan approximately 50% of the time and eventually solved 96% of their problems.

The qualitative differences in approach manifested by children of the various age groups and adults suggest that the system used in attempting to solve a problem might reflect different stages of conceptual thinking. An analysis of the demands inherent in the use of each strategy suggests that the ability to employ a particular strategy may be a function of the cognitive competence of the problem solver. The use of dimension checking seems to imply a knowledge of class inclusion concepts. Therefore, it is not surprising that this strategy was not commonly observed in kindergarten children. The use of focusing requires that the problem solver begins the problem with a more sophisticated long range plan. Initially, he considers each member of the total set of hypotheses the potential solution and remembers all of them. The set of potential solutions is systematically reduced by successively eliminating all hypotheses associated with negative feedback. The successful use of focusing appears to require a process of problem solving analysis similar to that postulated as implicated in the change in the relevant to irrelevant ratio observed in the 11 to 13 age range in the selective attention literature: i.e., the ability to accurately assess a task situation in its entirety and inaugurate a plan based on an analysis of the total set of inherent possibilities for solution. Therefore, the finding that virtually none of the problems of children from grades two to six showed this strategy is also not surprising. A focusing approach seems to call for the intellectual resources implicit in formal operational thinking (Flavell, 1963).

The present study is based on theoretical assumptions similar to those postulated by Flavell (1971) in his use of the term 'applied cognition' to refer to the processes involved in problem solving, learning and memory. It is assumed that children at different developmental stages have various cognitive competencies that reflect different underlying cognitive structural organizations. The choice of age groups of subjects for this study was based upon that assumption. The Piagetian literature (Piaget & Inhelder, 1969) suggests that third-grade children (roughly eight to nine) are most likely in the concrete operational stage while eighth-grade children (roughly thirteen to fourteen) should be capable of formal operational thinking. A second consideration in the choice of these age groups was that they have been repeatedly sampled in the selective attention literature and the shift of attention to relevant information has been consistently observed in the age range of the older group.

The experimental plan called for the analysis of problem solving behavior of the aforementioned age groups in four different stimulus situations. Observations of behavior across ages under the different conditions were expected to lead to inferences concerning the issues discussed above.

Stimulus situations differed in terms of whether or not all the dimensions in the stimulus complex were potentially the solution and whether or not a memory aid was provided.

In accord with the theoretical suppositions underlying the choice of age groups, it was expected that children in each group would manifest strategies appropriate to their assumed cognitive level in a situation in which no irrelevant information was included in the stimulus complex. Therefore, it was hypothesized that eighth-grade children would manifest focusing and dimension checking strategies in proportions approximate to the adult population in the Gholson et al. (1972) study. Third-grade children were expected to follow the pattern observed in elementary school children in the Gholson et al. (1972, 1973) studies; i.e., manifest dimension checking on the majority of the problems. The results of the Gholson et al. (1972, 1973) studies also led to the expectation that the older children would successfully complete a larger proportion of their strategic plan and solve more problems than the younger children.

Eimas (1970) has suggested that age differences in retaining information are an important factor in age related differences found in performance in concept identification problems. In this task, one way these differences would be manifested is in the tendency to repeat previously disconfirmed hypotheses. Neimark, Slotnich and Ulrich (1971) have noted that in recall tasks children, in contrast to adults, had difficulty in remembering which items they had previously recalled. The authors suggest that this is a result of relatively inefficient mnemonic encoding. In line with this, pilot data indicated that children in the third-grade, in contrast to the eighth-grade, had difficulty in remembering the hypotheses they had previously sampled. Thus, it was expected that younger children would resample a larger proportion of previously disconfirmed hypotheses than older children.

It was expected that children in these two age groups would also be dissimilar in their behavior in the task when irrelevant dimensions were included in the stimulus complex. In concordance with the assumption that cognitive level influences the behavior of children in a task situation which includes irrelevant information, it was assumed that eighth-grade children, but not third-grade children, would assess the situation and attempt to encode the relevant set of hypotheses in some mnemonic fashion. Consequently, it was hypothesized that the third-grade children would manifest a larger proportion of irrelevant hypotheses than the eighth-grade children.

The assumption that the older, but not the younger, children would plan their problem solving strategies based upon a thorough analysis of the task situation suggests that the eighth-grade children, but not the third-grade children, would alter their strategic approach to the problem when irrelevant information was included in the stimulus complex. It was expected that they would be more likely to initiate a plan which involved less 'cognitive strain' in this situation than in a situation in which no irrelevant information was present. Therefore, it was expected that a focusing strategy would be less evident in this age group in the former situation than in the latter. Consequently, there would be less difference in strategic approach between the two age groups when irrelevant information was present than when it was absent.

The presence of irrelevant information in the stimulus complex was expected to result in response disorganization among children of both ages due to the difficulty of relating feedback to relevant information only. Thus, it was expected that children of both age groups would show less efficiency in testing hypotheses and solve fewer problems than their peers exposed to a stimulus complex which did not include irrelevant information. Nevertheless, the older children were expected to show more efficiency than the younger and solve more problems.

Providing a memory aid that lists the total set of dimensions in the stimulus complex in a situation in which the solution might be any of them provided a test of an hypothesis suggested by Ingalls and Dickerson (1969) and Gumer and Levine (1971). Data from these studies indicate that subjects of all ages sometimes omit a dimension from the set of hypotheses considered. If the solution dimension is omitted the problem is, of course, unsolvable. It was also thought that the omission of a dimension might account for some of the 'random' patterns observed in the Gholson et al. (1972, 1973) studies. In addition, it was considered possible that the manifestation of previously disconfirmed hypotheses might be due to the child's sampling from a reduced set. Therefore, the provision of a memory aid when all the stimulus information was relevant was expected to result in eighth-grade children manifesting a more perfect version of particular hypothesis sampling systems than were manifested by their peers who were not provided with memory aids; i.e., more focusing and dimension checking and fewer random sequences. They were expected to manifest fewer previously disconfirmed hypotheses and to solve more problems. It was thought that third-grade children might not simultaneously attend to a memory aid and stimulus and feedback information. In that case, there would be no difference in performance between the third-grade groups exposed to relevant information only regardless of whether a memory aid was provided. However, if they did attend to the memory aid, it was expected to improve third-grade performance in the same way it improved eighth-grade performance.

It is difficult to tell from the Hagen et al. (1965, 1967, 1968, 1969, 1970, 1973) studies whether the young child's propensity to recall approximately the same proportion of relevant and irrelevant information is a function of difficulty in screening out irrelevant information, or a function of difficulty in encoding relevant information, or both. Providing a memory aid in the form of a list of only those dimensions from which the solution may be drawn in a problem solving situation in which other dimensions (irrelevant information) are included in the stimulus complex involved an attempt to disentangle the processes involved in excluding irrelevant information from attention and encoding relevant information for recall. This situation was conceived as furnishing an opportunity to explore the possibility that the exclusion of information from attention and the encoding of information for recall involve separate processes which may follow different developmental patterns.

If the younger children's difficulties in selective attention tasks are solely a function of inadequate or inappropriate encoding strategies as Hagen (1972) has hypothesized, the performance of third-grade children in a situation which includes both relevant and irrelevant information and a memory aid listing the relevant dimensions should be similar to that of their peers exposed to a stimulus complex without irrelevant information and provided with a memory aid. However, if the exclusion of irrelevant information from processing in itself is difficult, then their performance should be closer to that of their peers exposed to the same stimulus complex without the provision of a memory aid.

The finding that irrelevant information interfered with recall of relevant information in seventh-grade children but not college students (Hagen, 1967; Hagen et al., 1970) suggests that the presence of irrelevant information may interfere with encoding of relevant information even in older children. If the presence of irrelevant information has such an effect on the functioning of eighth-grade children then the provision of a memory aid should not improve performance in this situation. However, it was hypothesized that the older children's ability to assess the goals of the task coupled with a repertoire of strategic approaches would enable them to control their attentional response to irrelevant information. Therefore, it was predicted that the performance of eighth-grade children provided with a memory aid in this situation would be similar to that of their peers who were provided with a memory aid in a situation in which only relevant information was present.

In summary, the intention of this study was to investigate the problem solving behavior of third- and eighth-grade children presented with discrimination learning problems. The primary focus of the research was a comparison of hypothesis sampling systems manifested by children in these two age groups when all four dimensions included in a stimulus complex were potentially the solution (all relevant information) and when the stimulus complex also included four extraneous dimensions (irrelevant information) which could not be the solution. In addition, in each of these situations, the effect of providing a memory aid listing the four potential solution (relevant) dimensions of a particular problem was examined.

Method

Subjects: The subjects were 96 third-grade (mean CA = 8:10) and 96 eighth-grade (mean CA 13:10) children attending elementary and junior high schools in a middle class suburb of New York City. All third-grade children with parental permission were tested in one elementary school. Ten randomly chosen third-grade children from another elementary school in the district completed the third-grade sample. Junior high school students volunteered their free time. All who volunteered with parental permission were tested. Cells were filled in balanced order with the following restrictions: there were an equal number of boys and girls in each cell (12 boys and 12 girls); children assigned to the various cells were roughly matched on both age and the number of preliminary problems solved (solved one or two problems; solved three problems; solved four or five problems).

Design: There were foue experimental conditions.

Four-Dimension-No-Cue: Four-dimensional stimuli with no memory aid.

Eight-Dimension-No-Cue: Eight-dimensional stimuli with no memory aid.

Four-Dimension-Cue: Four dimensional stimuli with a memory aid.

Eight-Dimension-Cue: Eight-dimensional stimuli with a memory aid.

This resulted in a 2 x 2 x 2 design: two levels of age (third vs eighth grade) by two levels of information(four dimension (relevant only) vs eight dimension (irrelevant added)) by two levels of cue (memory aid vs no aid).

Materials: Eight four-dimensional bivalued problems modeled on Gholson et al. (1972) stimuli were used in two conditions. The same number of eight-dimensional bivalued problems constructed and organized according to the same principles were the stimuli in the other two conditions. The eight dimensions were shape (square and circle), letter (A and F; J and W;

H and Z; X and R; T and L), size (three and one inch), color (brown and green; purple and green; blue and red), line position (above and below the configuration), line quality (dashed and solid), border number (one and two), density (letter filled in and empty). The four-dimensional problems incorporated all of the dimensions used in the eight-dimensional problems. For each four-dimensional problem, there was a yoked eight-dimensional problem in which the same four dimensions were potentially the solution (relevant). In addition, four dimensions were included in the configuration that could not be the solution (irrelevant). Each problem was presented on twelve five x eight cards connected with loose leaf rings.

Stimulus Sequence: The stimuli were arranged with the special restrictions for orthogonality specified by Levine (1966). Each value of each dimension was counterbalanced with each value of every other dimension. The four stimulus pairs in a set were used to construct four sequential trials (i.e., trials 1-4, 5-8, 9-12). The four relevant stimulus pairs were arranged so that each of the set of hypotheses yielded a pattern of three responses to one side and one to the other during any four trials (e.g., right, right, right, left; left, right, left, left; left, right, right, right; right, right, left, right). Irrelevant hypotheses conformed to patterns of two or four responses per side (e.g., left, left, left, left; left, right, left, right; left, left, right, right; left, right, right, left). The problems were constructed so that the solution hypothesis was logically defined in any three consecutive feedback trials (see Appendix B for a four dimension example).

All trials were feedback trials. Hypotheses probes preceded each feedback trial. A verbal probe was the method of obtaining hypothesis information from trial to trial. The decision to use a verbal probe was based on Karpf and Levine's (1971) finding that verbal probes yield the same results as blank trials in adults. This was substantiated by Phillips (1973) dissertation comparing blank-trial probes, verbal probes and a no probe condition among groups of second- and sixth-grade children. Her data suggested that this was a valid method for obtaining successive hypotheses from elementary school children. Since there were practical disadvantages in the use of the blank-trial probe (i.e., loss of data for system analysis if there are any inconsistencies in the first three blank-trial probes and intervening feedback trials), the verbal probe was used. Feedback followed each trial. The sequence was: the subject stated his hypothesis while the stimulus card was in front of him; he chose one of the pair of stimulus configurations presented by pointing to his choice; he was given verbal feedback as to the correctness of his choice (e.g.,"yes, that's right" or "no, that one isn't right, this one is" pointing to the stimulus configuration that included the solution.) The stimulus configurations remained in front of the subject for about three seconds following feedback. This procedure was continued through twelve such trials.

Feedback information was preprogrammed so that each subject received the same feedback sequences during the first three trials of the various problems. The feedback patterns used were positive feedback on the first two trials and negative on the third trial (++-); negative feedback on each of the first three trials (---); positive feedback on the first trial and negative feedback on the ensuing two trials (+--); negative feedback on the first trial, positive feedback on the second trial and negative feedback on the third trial (-+-). Each pattern was assigned to a given problem (two problems per pattern). The solution that was logically specified following the third trial, which always involved negative feedback, was the basis for further feedback information.

Procedure: Each subject was seen individually in an office in his own school. The experimenter and the child sat side by side at a desk. Preliminary training, consisting of five practice problems, was the same for all children. The first two problems were four and eight dimension, respectively. During these problems, the children were instructed in the nature of the task; i.e., that one value was the solution; that there was a defined finite set of potential solutions; the method of stimulus and feedback presentation. The next three preliminary problems were used to instruct the subject in the method of verbally stating hypotheses. He was taught to verbalize his hypothesis before each choice response without prompting. He was asked to state an hypothesis prior to his first choice response. He was told that this was optional but verbal statements after the first feedback trial were mandatory. He was instructed that feedback information referred to responses to the stimulus complex only. The third problem was four-dimensional and the last two were eight-dimensional.

Following preliminary training, each child was presented with a series of eight experimental problems. Before each problem, the set of possible solution hypotheses was verbally described by the experimenter while she pointed them out in the stimulus configurations: e.g., "In this problem, the right answer might be 'red' or 'blue' or 'small' or 'big' or 'line up' or 'line down' or 'circle' or 'square'". The child was then asked to repeat them and point to each as it was named. He was corrected if he made a mistake. There was no further instructions in the no memory aid conditions (see Appendix C for exact instructions).

In the memory aid condition, after the experimenter had labeled the hypotheses and the child had repeated them perfectly, a cue card (5 x 8 inches) listing the entire set was placed before him above the booklet that contained the problem. He was asked to read the list in order to assure that he could read. He was told that he could use the card during the problem as a reminder of the set of possible answers (see Appendix D for copy of sample cue card).

Following the final problem each subject was asked how he solved the problems.

Results and Preliminary Discussion

The following analyses were based on the 96 children of each grade level (12 males and 12 females in each of eight cells) who began the experiment. Experimental error on the fourth feedback trial resulted in the loss of data on 11 problems (one from each of the following subjects: one male and one female-third-grade, Four-Dimension-No-Cue, one male and two females-third-grade, Four-Dimension-Cue, two males-third-grade, Eight-Dimension-No-Cue, one female-eighth-grade, Four-Dimension-No Cue, one female-eighth-grade, Four-Dimension-Cue and two problems for one female-third-grade, Eight-Dimension-Cue). The analyses were based on proportion of each child's responses normalized by the Freeman-Tukey arcsin transformation for binomial proportions (Freeman & Tukey, 1950). Strategy data for six of the problems were available and are included in the strategy analyses (one male and one female-third-grade, Four-Dimension-Cue, one male, third-grade, Eight-Dimension-No-Cue, one female-eighth-grade, Four-Dimension-Cue and the two problems of the female in the third-grade, Eight-Dimension-Cue).

Response Consistency: In view of the fact that the verbalized hypotheses are the primary data used in the analyses of problem solving behavior in this investigation, it is important to note that across both ages and over all conditions, the response choice of the child was always predicted by his verbalized hypothesis; that is, the verbalized hypothesis corresponded to one of the four values in the stimulus complex chosen directly following verbalization 100% of the time.

Effects of Sex: The examination of the results was begun with statistical analyses of the effects of sex on two overall measures of problem solving efficiency.

Probability of solution: A problem was designated as solved if the solution hypothesis was verbalized on the ninth trial and continued to be held throughout the remaining trials (9-12). For each child, the ratio of the total problems solved to those completed was computed. A 2 x 2 analysis of variance (Age x Sex) was performed and yielded a reliable main effect for age. $F(1,188)=46.355$. $p < .05$. The main effect for sex and the age x sex interaction were not significant ($F=1.33$ and $F < 1$ respectively). The procedure throughout this paper will be to report only one level of significance ($p < .05$).

Probability of retaining an hypothesis given positive feedback: One basic assumption of hypothesis testing theory is that the subject retains an hypothesis when feedback is positive (Gholson et al., 1969; Levine, 1966, 1969; Trabasso and Bower, 1968). The proportion of trials in which an hypothesis was retained when followed by positive feedback ($P(H_i=H_{i-1} \mid F_i=+)$ was calculated for each subject. A 2 x 2 analysis of variance yielded a reliable main effect of age, $F(1,188)=27.74$, but no significant sex or sex x age interaction effects ($F < 1$, in both cases). In view of these preliminary findings, the remaining analyses were performed on data that were collapsed over sex.

Solution Data:

Probability of Solution: A 2 x 2 x 2 (Age x Information x Cue) analysis of variance was performed on the probability of solution. There were significant main effects of age ($F(1,184)=65.23$), information ($F=57.365$) and cue ($F=18.76$) but no interaction effects were significant (see Appendix E for all tables listing analyses of variance results). The mean proportion of problems solved per experimental condition is presented in Table 1. Since the main purpose of this study

Table 1

Mean Proportion of Problems Solved for Each Experimental Group

Condition	Grade	
	Third	Eighth
Four-Dimension-No-Cue	.679	.922
Four-Dimension-Cue	.804	.937
Eight-Dimension-No-Cue	.532	.688
Eight-Dimension-Cue	.658	.844

was to compare performance under different experimental conditions, comparisons between the various groups were especially relevant to the experimental questions. An appropriate technique for these comparisons is the Tukey b procedure applied to ordered means (Petrinovich & Hardyk, 1969). It is to be noted that this is a more conservative estimate than the F test and leads to fewer significant differences.

With this procedure, each eighth-grade group, with the exception of Eight-Dimension-No-Cue, solved a significantly larger proportion of problems than the comparable third-grade group. A comparison of information conditions in which no cue was provided (Four-Dimension-No-Cue vs Eight-Dimension-No-Cue) indicated that both age groups solved a significantly larger proportion of problems in the four dimension condition than in the eight dimension condition. The presence of a cue list appeared, however, to have differential effects depending on age and information level. The proportion of problem solved by the third-grade children in the Four-Dimension-Cue condition was significantly larger than the proportion solved by their peers in both the Four-Dimension-No-Cue and Eight-Dimension-Cue conditions. But there was no significant difference in the proportion of problems solved by the two third-grade Eight-Dimension groups (Cue, No Cue). Among the eighth-grade groups, there was no significant difference in the proportion of problems solved by the Four-Dimension groups (Cue, vs No Cue) nor between the groups provided with a cue (Four-Dimension-Cue vs Eight-Dimension-Cue). But the eighth-grade group provided with a cue with eight-dimensional stimuli solved a significantly larger proportion of problems than the eighth-grade group exposed to the same stimuli without a cue list.

Thus, the third-grade group provided with a cue with stimuli that contained no irrelevant dimensions solved significantly more problems than any other third-grade group. Also, exposure to only relevant information rather than the combined relevant-irrelevant complex resulted in a significantly larger proportion of solved problems in this age group. Providing a cue did not affect third-grade problem solution when the stimuli contained irrelevant information. On the other hand, the proportion of problems solved by the eighth-grade children was not significantly different when no irrelevant information was included in the stimulus complex with or without a cue or when the combined relevant-irrelevant complex was presented with a cue. But presenting the relevant-irrelevant complex without a cue significantly lowered the proportion of problems solved.

Errors in unsolved problems: It seems appropriate at this point to examine the protocols of all unsolved problems to determine the errors that interfered with solution. Three types of error were identified:

1. The solution dimension was never sampled.
2. The value opposite to the correct hypothesis of the solution dimension was the only one manifested on that dimension as either the initial hypothesis (H_0) which was stated prior to first feedback or in the body of the problem (i.e., hypotheses 1 through 8).
3. The solution hypothesis was sampled and rejected despite positive feedback.

An examination of Table 2 reveals that the most frequent error that prevented solution among the third-grade children was neglecting to sample the solution dimension (accounting for 70% of the failures in this age group). Sampling only the value opposite the correct hypothesis, either within the body of the problem or as the first hypothesis prior to feedback was the second most frequent error (accounting for 25% of the failures). On the remaining 5%, the solution hypothesis was sampled and rejected despite positive feedback.

As Table 2 indicated, the third-grade pattern of errors in unsolved problems remained constant across conditions, whereas the proportional distribution of errors in eighth-grade groups appeared to have been affected by the experimental conditions. The presence of a cue list increased the proportion of problems in which only the value opposite the correct hypothesis of the solution dimension was sampled. This is particularly apparent in the Four-Dimension-Cue group in which the usual pattern was reversed in that only the value opposite the correct hypothesis of the solution dimension was sampled on 67% of the unsolved problems either as H_0 or following feedback. These data suggest that the presence of a cue list increased the likelihood that the eighth-grade children included all of the relevant dimensions in their sample set but it was at times also distracting in that it increased the probability of their sampling the wrong hypothesis on the solution dimension.

However, there are wide differences in the number of problems on which these proportions are based and these proportions are completely interdependent. Since the eighth-grade children solved more problems than the third-grade children in every condition (see Note on Table 2), there were fewer problems represented in each of the eighth-grade groups. These proportions, therefore, may be somewhat misleading. Bearing this in mind, the failure to sample the solution dimension accounted for 66% of the unsolved eighth-grade problems while sampling only the value opposite the correct hypothesis on the solution dimension accounted for 30% of the unsolved problems. The remainder (4%) showed the rejection of the solution hypothesis despite positive feedback.

Table 2

Percentage of Errors in Unsolved Problems in which either the Solution Dimension was Missing, only the Dimensional Opposite Hypothesis to Solution was Verbalized or the Solution was Hypothesized and not Maintained

	Condition	Grade	
		Third	Eighth
Solution dimension missing	Four-Dimension-No-Cue	69	60
	Four-Dimension-Cue	73	33
	Eight-Dimension-No-Cue	73	75
	Eight-Dimension-Cue	65	63
	Total	70	66
Only dimensional opposite hypothesis verbalized	Four-Dimension-No-Cue	25	27
	Four-Dimension-Cue	24	67
	Eight-Dimension-No-Cue	25	27
	Eight-Dimension-Cue	23	33
	Total	25	30
Solution hypothesized and not maintained	Four-Dimension-No-Cue	6	13
	Four-Dimension-Cue	3	0
	Eight-Dimension-No-Cue	2	2
	Eight-Dimension-Cue	12	3
	Total	5	4

Note. Number of unsolved problems per group

	Third	Eighth
Four-Dimension-No-Cue	61	15
Four-Dimension-Cue	37	12
Eight-Dimension-No-Cue	89	60
Eight-Dimension-Cue	65	30

Thus, in the majority of unsolved problems among children of both age groups, the solution dimension was omitted. This error has also been observed by Levine et al. (1966) and Gumer & Levine (1971) in adult protocols using blank-trial hypotheses probes.

In an attempt to determine the cause of this error, an examination was made of the number of dimensions sampled during those problems in which neither value of the solution dimension was tested. As Table 3 indicates, the third-grade children tested only two dimensions in 29% of the four-dimensional problems when no cue was present. This was reduced to 11% when the cue was present. The eighth-grade children tested hypotheses from three dimensions during all four-dimensional problems in which hypotheses from the solution dimension were omitted. This suggests that failure to solve in the majority of cases is due to difficulty in processing information in relation to four dimensions and that the younger children are likely to have more difficulty in this regard than the older children. But this is contradicted by the data from the eight-dimensional problems. When no cue list was provided with eight-dimensional stimuli, the third-grade children tested hypotheses from four to six dimensions (relevant and irrelevant) during 54% of those problems in which the solution dimension was omitted. The percentage for the eighth-grade children was 69%. Providing a cue with these stimuli reduced the proportion in both age groups, particularly the eighth-grade group. The sampling by both age groups of four or more dimensions in over 50% of the eight-dimensional problems in which the solution dimension was omitted does <u>not</u> support the hypothesis that failure to sample the solution dimension is a function of difficulty in processing information in relation to a large number of dimensions. When all dimensions

Table 3

Percentage of Unsolved Problems Missing the Solution Dimension in Which Hypothesis Testing Included 1 to 6 Dimensions

Group	Dimensions						Number of Problems
	1	2	3	4	5	6	
Third grade							
Four-Dimension - No-Cue	0	29	71	-	-	-	42
Four-Dimension - Cue	0	11	89	-	-	-	27
Eight-Dimension - No-Cue	0	3	43		54		65
Eight-Dimension - Cue	4	11	50		35		42
Eighth grade							
Four-Dimension - No-Cue	0	0	100	-	-	-	9
Four-Dimension - Cue	0	0	100				4
Eight-Dimension - No-Cue	0	4	27		69		45
Eight-Dimension - Cue	0	11	63		26		19

were relevant, providing a cue list increased the probability that all dimensions would be sampled in both age groups. Although children in both age groups sampled more dimensions when irrelevant dimensions were included in the stimulus configuration, they also were more likely to omit the solution dimension. Providing a cue with this stimulus configuration reduced the proportion of problems in which four or more dimensions were sampled but also reduced the proportion of problems in which the solution dimension was omitted in both age groups. Thus, it appears that failure to sample the solution dimension may be due to the subject's failure to properly define the relevant set of hypotheses i.e., to include all relevant dimensions and exclude all irrelevant dimensions. The data from this study suggest that third-grade children are less likely to properly define the sampling set than eighth-grade children, regardless of experimental conditions.

Problem order effects: In view of the experience with five practice problems exemplifying the structure and procedure of the task, it was not expected that the probability of solution would increase across the eight experimental problems (i.e., that learning-to-learn would occur). The emphasis in a procedure of this kind is on the application of methods of problem solving that are presumed to be in the repertoire of the subject, rather than the learning of new problem solving techniques within the experimental session. Nevertheless, it is of interest to examine this question statistically. Based on the assumption that there is a constant probability of solving each problem, there should be no difference in frequency of solution between the first and second blocks of problems (problems 1-4 vs problems 5-8). A χ^2 (chi square) binomial distribution with three degrees of freedom was devised for each cell separately. Theoretical probabilities were based on the proportion of problems solved within each group and p and q theoretical probabilities were based on the frequency data (personal communication, Sol Weinstock).

The overall 2 x 2 χ^2 (24)=29.72($\underline{p}$=.20) was not significant. The only significant increase in the probability of solution for the second block over the first block took place in the third-grade Four-Dimension-Cue condition (χ^2(3)=7.84, $\underline{p} < .05$). The only other analysis that approached significance was in the eighth-grade, Eight-Dimension-Cue condition (χ^2(3)=6.18, $\underline{p}$ .10 $> \underline{p} <$ .20) (See Appendix F for table listing the Chi Squares for each group and the overall Chi Square for the eight groups and the frequencies across problems for each of the experimental groups).

Hypothesis Testing Efficiency Data:

The probability of retaining an hypothesis following positive feedback: The assumption of Hypothesis theory that the subject retains a confirmed hypothesis was examined by looking at each consecutive pair of hypotheses when the intervening feedback was positive. The results of a 2 x 2 x 2 analysis of variance revealed that the main effects of both age and information were significant ($\underline{F}$(1,184)=29.3 and 9.0 respectively). The only interaction that approached significance was the age x cue ($\underline{F}$=2.8,$\underline{p}$ .10 $>$ p $>$ .05). The probability of retaining a confirmed hypothesis ($P(H_i - H_{i-1} \mid F_i = +)$ was over .90 in each condition. Within each condition, the eighth-grade children showed a greater probability of maintaining an hypothesis following positive feedback than the third-grade children. Overall, both age groups showed more efficiency in this regard with four-dimentional stimuli than eight-dimensional stimuli. With the more conservative Tukey $\underline{b}$ test, the only significant difference between experimentally relevant probabilities occurred between the two age groups at the two information levels under the No Cue conditions (third-vs eighth-grade, Four-Dimension-No-Cue, Eight-Dimension-No-Cue.)

The aforementioned analysis included data across the 12 trials of every problem. Including data from the criterion run may inflate the proportion of confirmed hypotheses that are maintained. If so, it also gives differential weight to the proportions within each experimental group as functions of the percentage of problems solved and the trial per problem on which the criterion run began. Further, since the verbalization of H_0 was optional, the proportion of problems on which it was offered differed across experimental groups. H_0 was more likely to be observed in third-grade data than eighth-grade data (81% vs 62% of the problems). In addition, when H_0 was verbalized, there was a tendency for children in both age groups to shift to a new hypothesis following positive feedback on Trial 1. This shift occurred on 23% of the third-grade and 26% of the eighth-grade problems in which Trial 1 feedback was positive. Therefore, another measure of the effects of positive feedback excluding both criterion run data and H_0 data was computed. The results of the 2 x 2 x 2 analysis of variance performed on these data reinforced the effects found when all data were included in the analysis. Again, only age ($F(1,184)=11.8$) and information ($F=7.8$) main effects were significant. But, in this analysis, the age x cue interaction was also significant ($F=3.9$) and the cue x information interaction approached significance ($F=2.7$, p $.10 > p > .05$). The only comparison that reached statistical significance with the Tukey b test was in the eighth-grade between information levels when a cue was provided (Four-Dimension-Cue vs Eight-Dimension-Cue). Figure 1 presents a

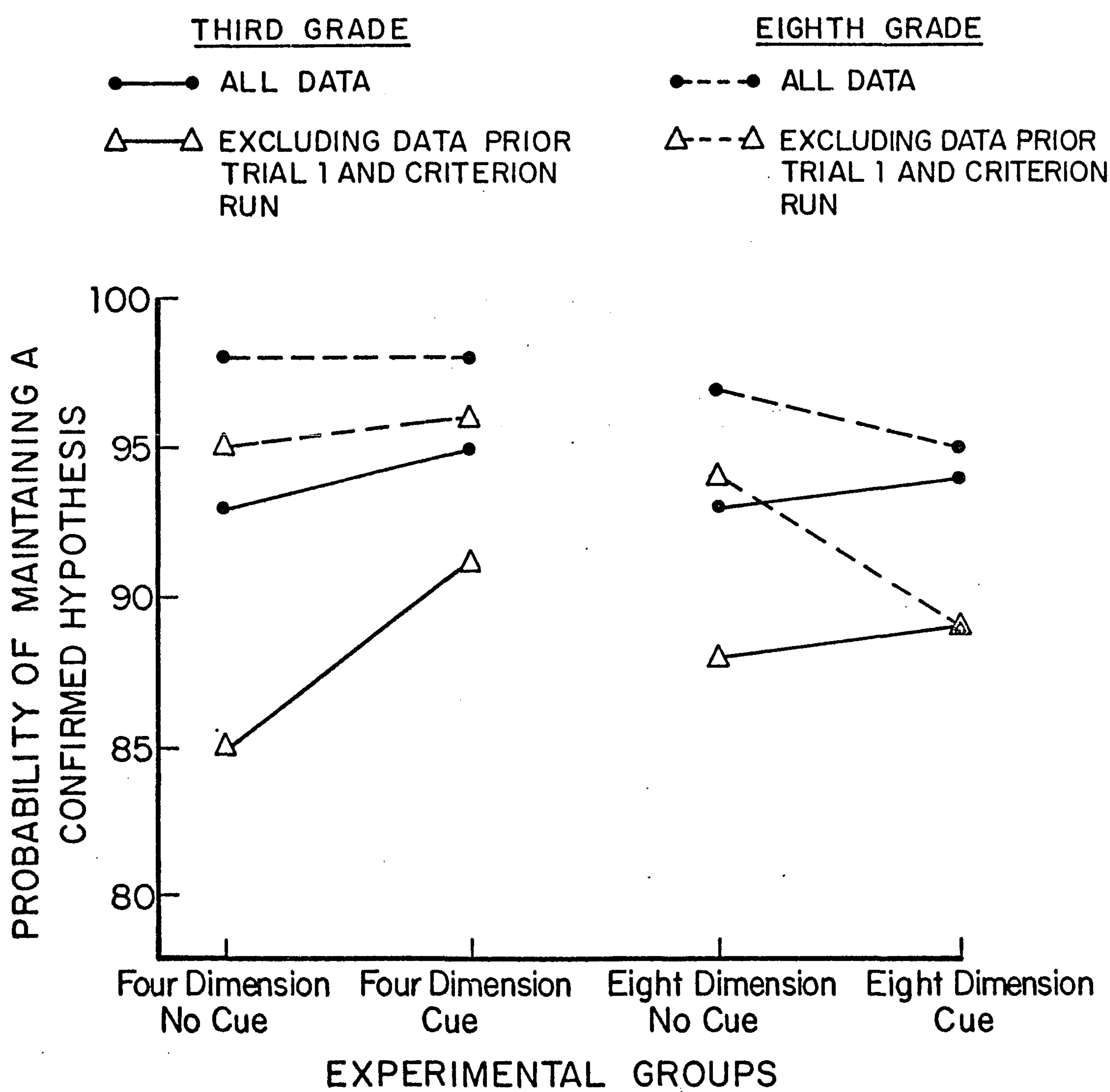


Fig. 1 Comparison of two measures of the probability of maintaining a confirmed hypothesis for each experimental group

comparison of the probabilities obtained when all data were included and when only data subsequent to Trial 1 through criterion were included. The probabilities in the second measure present a clearer picture of the differential effects of the cue list as a function of age and information. The third-grade children were more likely to retain an hypothesis following positive feedback when a cue was provided, particularly with four-dimensional stimuli. In contrast, the provision of a cue did not affect the tendency of eighth-grade children, exposed to four-dimensional stimuli, to retain an hypothesis following positive feedback. It did, however, decrease the probability that these children retained an hypothesis following positive feedback when the stimuli included irrelevant information. Thus the eighth-grade children were more likely than the third-grade children to retain an hypothesis following positive feedback in all conditions except Eight-Dimension-Cue.

But shifting from an hypothesis is not necessarily an erroneous response to positive feedback. It may be a different method of testing hypotheses (i.e., test 'black', hold it in abeyance, test 'square; given negative feedback, retest 'black'). Both Bruner et al. (1956) and Kemler (1972) have reported similar hypothesis sampling patterns with different kinds of problems. In the present study, children in both age groups not only showed a tendency to shift hypotheses following Trial 1 positive feedback but also following Trial 2 positive feedback (shift from first to second hypothesis- H_1 to H_2). Data on hypothesis shift following Trial 2 are included in both of the preceding analyses. Table 4 column 1 shows the proportion of <u>all</u> shifts subsequent to first feedback that immediately followed Trial 2 feedback and were totally consistent with all previous feedback (globally consistent). In each condition, a larger pro-

Table 4

Proportion of Hypothesis Shifts Following Positive Feedback that Immediately Followed Trial 2 Feedback and were Globally Consistent

Group	Proportion of all shifts	Proportion of H1 shifts	χ^2
Third grade			
Four-Dimension- No-Cue	.104	.438	2.81
Four-Dimension- Cue	.194	.700	6.34*
Eight-Dimension- No-Cue	.092	.429	2.33
Eight-Dimension -Cue	.036	.222	0.02
Total	.098		
Eighth grade			
Four-Dimension - No-Cue	.438	.778	7.41*
Four-Dimension - Cue	.600	.750	8.96*
Eight-Dimension - No-Cue	.267	.800	8.82*
Eight-Dimension -Cue	.298	.667	11.76*
Total	.352		

portion of eighth-grade than third-grade shifts fell into this category. In both age groups, the largest proportion of globally consistent shifts immediately following Trial 2 feedback were observed in the Four-Dimension-Cue condition.

Column 2 shows the proportion of the shifts following Trial 2 positive feedback that were globally consistent. Since a shift may be made to seven possible hypotheses (three locally consistent and four not locally consistent with Trial 2 feedback) and only one of these is consistent with both previous feedbacks, the probability that each of these proportions deviated from chance was tested by χ^2 (chi square) using the Yates correction for discontinuity.

Column 3 indicates that in each eighth-grade group, the proportion of such shifts that were globally consistent exceeded chance expectation. For the third-grade groups, this was only true in the Four-Dimension-Cue condition. Thus, it may be that the majority of eighth-grade shifts following Trial 2 are part of a systematic sequence of hypotheses sampling. Comparative third grade procedures were only observed when the stimuli contained no irrelevant dimensions and a cue list was provided (Four-Dimension-Cue).

There is another situation in which shifting hypotheses following positive feedback may be systematic. If, following positive feedback, the subject shifts to the solution hypothesis subsequent to Trial 4, and holds this hypothesis through criterion, the shift may represent a sudden recognition of the solution to the problem. In this case, the shift is also to the hypothesis consistent with all previous feedback. Shifts of this kind were, like those described above, more likely to be observed among eighth-graders than third-graders (over all conditions; eighth-grade, 25%, third-grade, 15%). Table 5, column 1 presents these data while column 2 shows the total proportion of all shifts that were consistent with all previous feedback (either shifts at H_2 or to solution). Over all conditions more than twice as many eighth-grade than third-grade shifts were globally consistent(eighth-grade 48%, third-grade 21%).

The probability of retaining an hypothesis following negative feedback: Shifting from an hypothesis following positive feedback may or may not represent an hypothesis testing error, but retaining an hypothesis following negative feedback ($P(H_i=H_{i-1} \mid F_i=-)$ always does, as the verbalized hypothesis always appeared in the stimulus choice that received negative feedback on such trials. The propensity to retain an hypothesis following negative feedback was investigated by examining consecutive pairs of hypotheses of each problem in which the intervening feedback was negative. A 2 x 2 x 2 analysis of variance revealed that age was the only significant main effect ($F(1,184)=12.19$) and there were no significant interactions.

As Table 6 indicates, the probability of maintaining a disconfirmed hypothesis was extremely small for eighth-grade children (less than .02 in any condition). Although the probability of committing this error was greater in the third-grade groups, children in this age group also showed a very small probability of maintaining disconfirmed hypotheses (.02 to .05).

The probability of generating a locally consistent hypothesis following negative feedback: When a choice response is followed by negative feedback, the child is given the information that none of the hypotheses in that configuration is the solution. Thus sampling any hypothesis from that configuration represents an inefficient mode of processing. Any hypothesis in the contrasting stimulus configuration is,

Table 5

Proportion of Hypothesis Shifts that were Globally Consistent

Group	Proportion Shifts Following Trial 4 to Solution	Total Proportion Shifts Globally Consistent
Third grade		
Four-Dimension-No-Cue	.137	.209
Four-Dimension-Cue	.201	.361
Eight-Dimension-No-Cue	.137	.200
Eight-Dimension-Cue	.149	.161
Total	.154	.219
Eighth grade		
Four-Dimension-No-Cue	.143	.500
Four-Dimension-Cue	.333	.667
Eight-Dimension-No-Cue	.150	.367
Eight-Dimension-Cue	.346	.489
Total	.250	.481

Table 6

Probability of Maintaining a Disconfirmed Hypothesis for Each Experimental Group

Condition	Third	Eighth
Four-Dimension-No-Cue	.051	.007
Four-Dimension-Cue	.022	.002
Eight-Dimension-No-Cue	.026	.017
Eight-Dimension-Cue	.040	.010

of course, consistent with the information given on that trial. Sampling a new hypothesis from the latter configuration is an indication of locally consistent sampling. This type of hypothesis testing efficiency was examined by computing the proportion of hypotheses per subject that were consistent with the preceding positive stimulus configuration following negative feedback. The results of a 2 x 2 x 2 analysis of variance showed significant main effects of age ($\underline{F}(1,184)=30.2$) and information ($\underline{F}=21.6$). Although the eighth-grade children showed a larger proportion of locally consistent hypotheses within each condition, the only age differences that were significant using the Tukey b method were in the Four-Dimension conditions (Cue and No-Cue). This method also revealed that the information main effects were primarily a function of the effect of irrelevant information on the functioning of eighth-grade children. The only significant information differences occurred between the eighth-grade groups, both with and without a cue (Four-Dimension-No-Cue vs Eight-Dimension-No-Cue and Four-Dimension-Cue vs Eight-Dimension-Cue). As Figure 2 reveals, the production of locally consistent hypotheses by eighth-grade children was sharply curtailed by the presence of irrelevant information whether or not a cue listing the relevant hypotheses was provided. On the other hand, there was not much difference in the production of locally consistent hypotheses by third-grade children regardless of whether irrelevant information was included in the stimulus complex. Although the third-grade children showed most local consistency in the Four-Dimension-Cue condition and least in the Eight-Dimension-Cue condition, the difference between the two groups did not reach statistical significance. Overall, children in both age groups were highly likely to generate locally consistent hypotheses regardless of experimental condition.

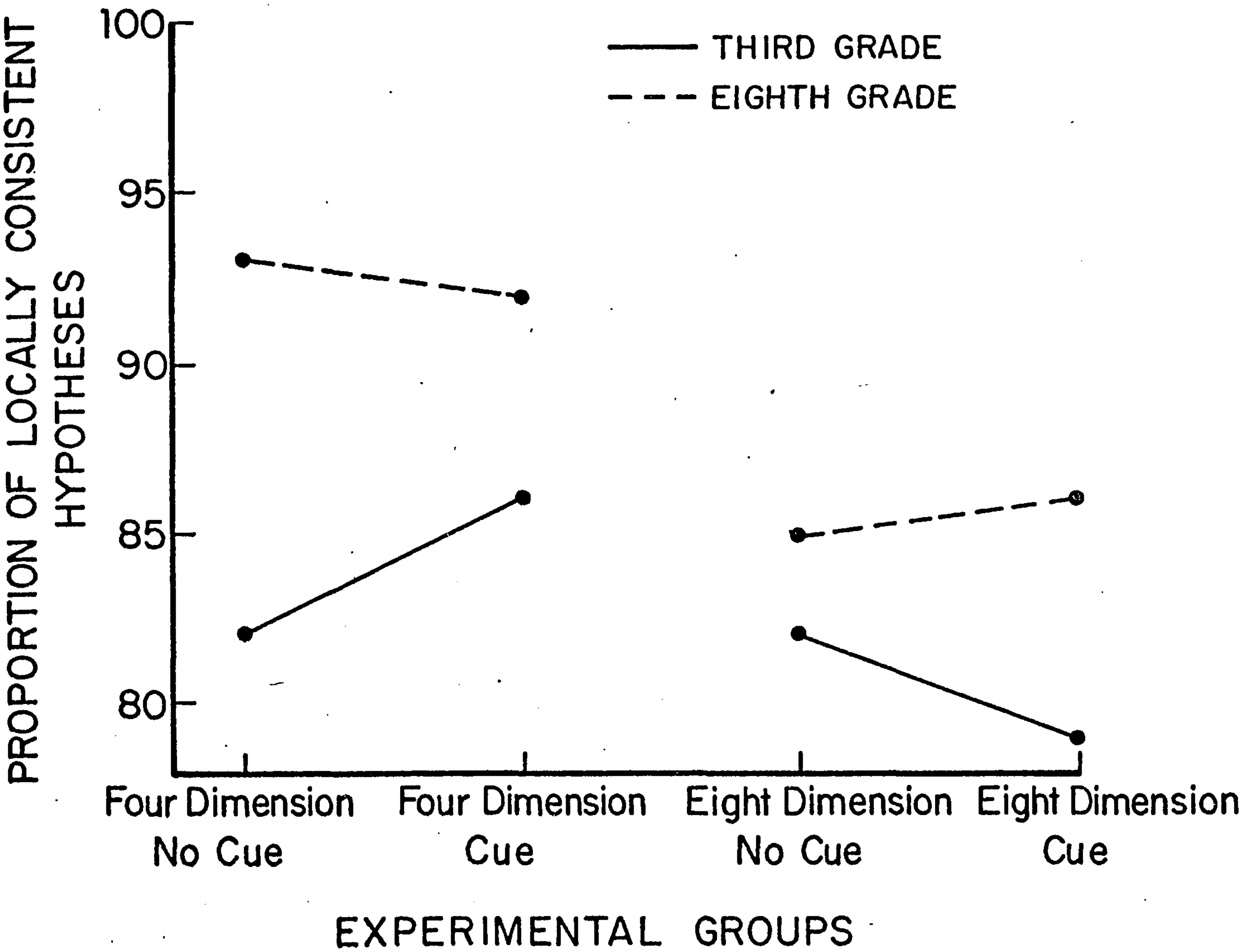


Fig. 2 Proportion of locally consistent hypotheses for each experimental group

The probability of manifesting a previously disconfirmed hypothesis: Another indication of hypothesis testing efficiency is evidenced when a new hypothesis is verbalized. It is, of course, very inefficient to retest (verbalize) an hypothesis that has been previously tried and disconfirmed. Retesting a previously disconfirmed hypothesis necessarily delays reaching solution. It is also likely that retesting a disconfirmed hypothesis represents some memory loss. For each problem, the proportion of disconfirmed hypotheses (i.e., hypotheses that had previously been manifested with ensuing choices receiving negative feedback) to new hypotheses (either following negative feedback or a shift following positive feedback) was computed. A 2 x 2 x 2 analysis of variance revealed significant effects of age ($F(1,184)=44.3$) and cue ($F=4.5$). The information effect approached significance ($F=2.9$, p $.10 < p < .05$) but none of the interactions did. Only the age comparisons in the Four-Dimension-No-Cue and Eight-Dimension-Cue conditions showed significant differences using the Tukey *b* method. Figure 3 presents these data for each group. Providing a cue with stimuli that included only relevant dimensions reduced the proportion of manifested hypotheses that had been previously disconfirmed among the third-grade children, whereas it had little effect on the eighth graders. Therefore, the age difference was only significant when no cue list was provided with four-dimensional stimuli. When irrelevant information was included in the configuration the eighth-grade children, but not the third-grade children, produced a larger proportion of previously disconfirmed hypotheses than they had when the configuration contained no irrelevant information. Consequently the mean difference between the age groups in this condition was not statistically significant. Although the provision of a cue list with

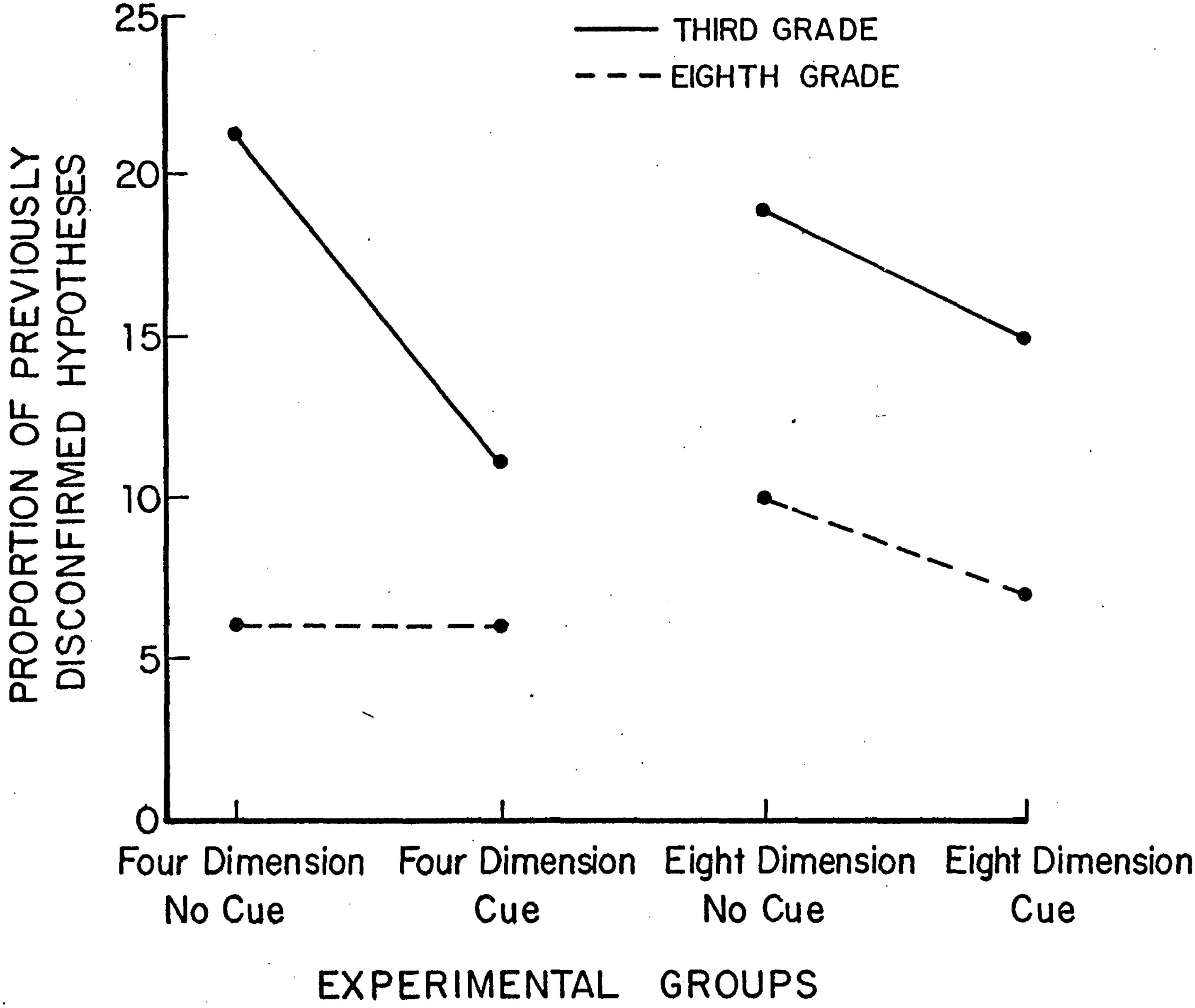


Fig. 3 Proportion of hypotheses previously disconfirmed for each experimental group

eight-dimensional stimuli lowered the proportion of previously disconfirmed hypotheses in both age groups, the difference between the two age groups was significant when the cue was present. Overall, there were no significant differences among the eighth-grade groups nor among the third-grade groups. Under all experimental conditions, the eighth-grade children were less likely than the third-grade children to retest an hypothesis that had been previously disconfirmed.

The probability of generating an irrelevant hypothesis: An indication of hypothesis testing efficiency that is central to the purpose of this study is the extent to which hypotheses drawn from irrelevant dimensions were observed. Only Eight-Dimension conditions are of interest here. Processing information in relation to any of the irrelevant hypotheses is inefficient as it necessarily delays solution. For each subject, the proportion of verbalized irrelevant hypotheses to all verbalized hypotheses was computed. A 2 x 2 analysis of variance was performed on these data. The main effects of age ($F(1,92)=8.2$) and cue ($F=57.5$) were significant. As Figure 4 indicates, irrelevant hypotheses were manifested among children of both age groups regardless of whether a cue list was provided. But fewer irrelevant hypotheses were manifested by the eighth-grade children than the third-grade children in both conditions. Also the provision of a cue list significantly reduced the proportion of irrelevant hypotheses manifested by children in both age groups. Another 2 x 2 analysis of variance was performed on the percentage of problems in these two conditions in which one or more irrelevant hypotheses was manifested. As Figure 5 shows, the results duplicated those obtained in the previous analysis.

It is clear from both figures that the two age groups differed

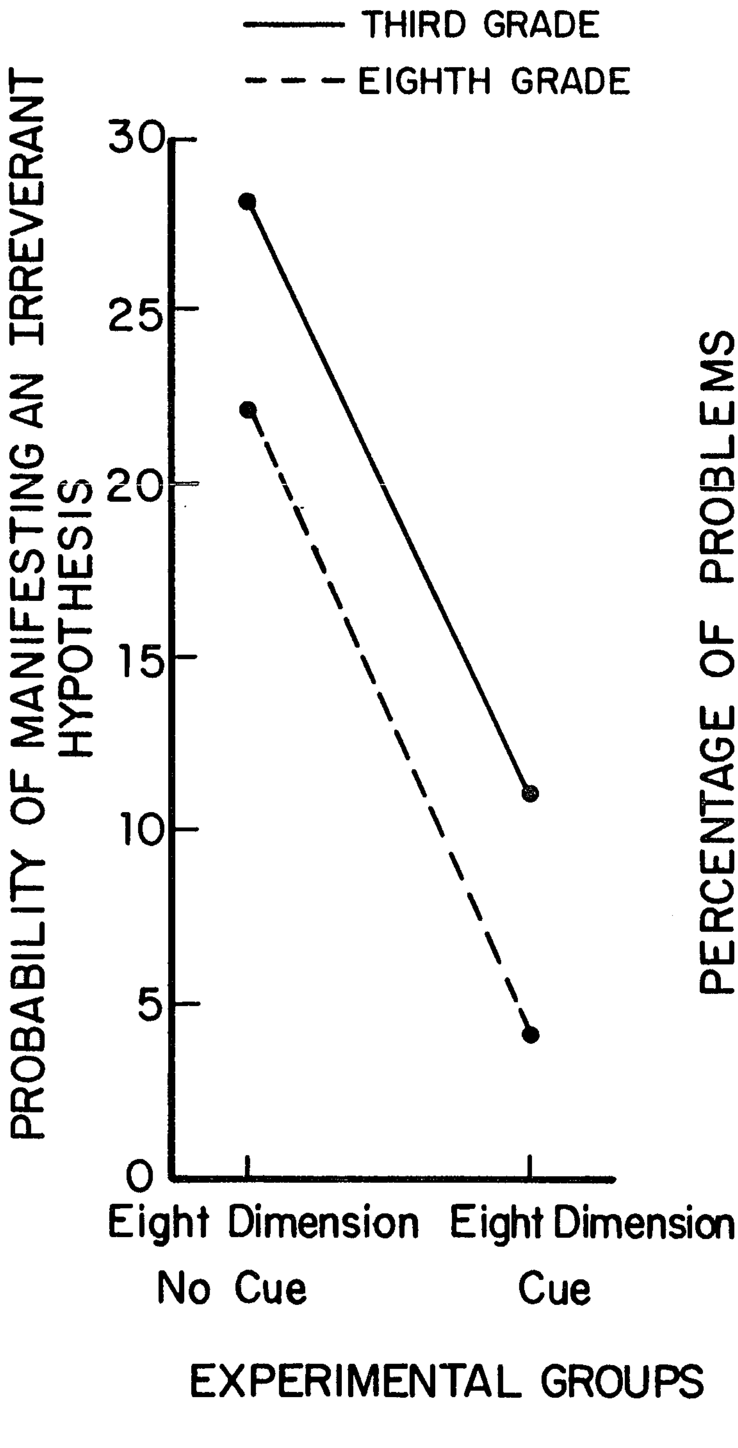


Fig. 4 Probability of manifesting an irrevelant hypothesis for each experimental group

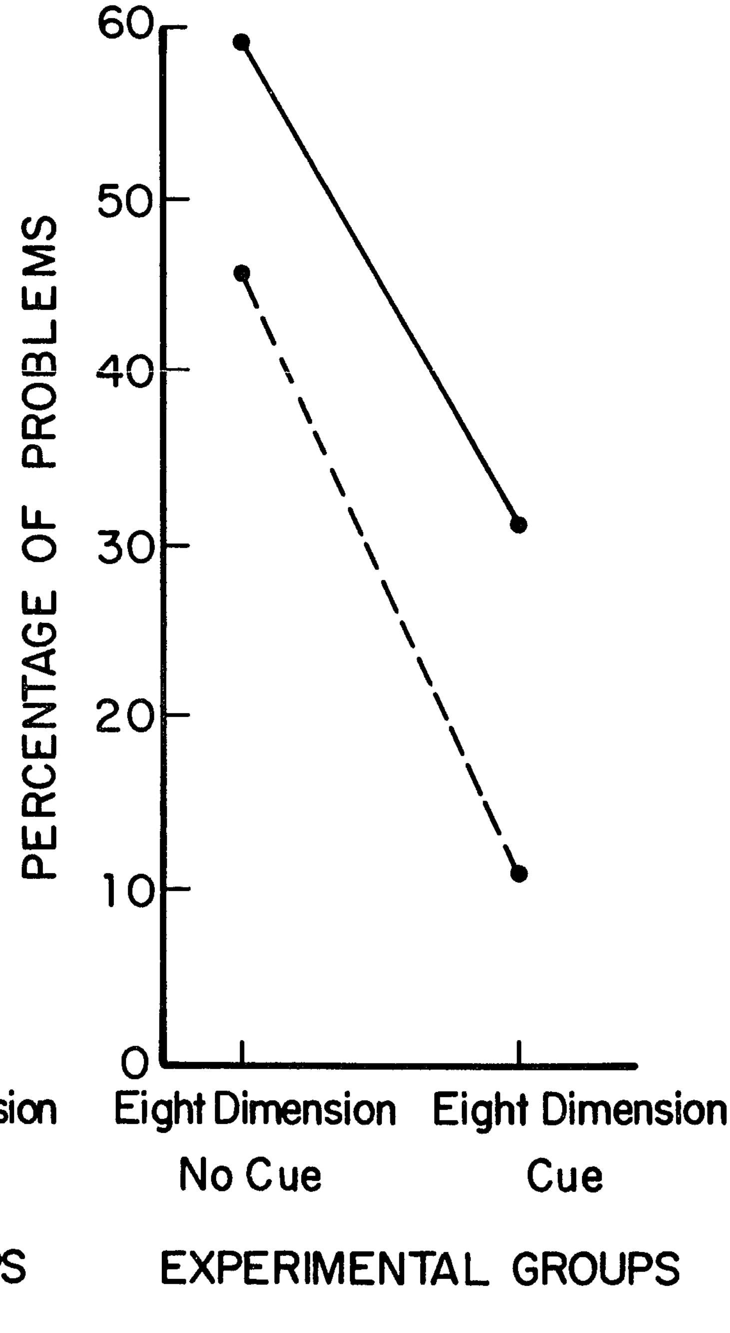


Fig. 5 Percentage of problems in which an irrevelant hypothesis was manifested for each experimental group

in both the proportion of problems in which an irrelevant hypothesis was observed and the proportion of irrelevant hypotheses sampled across problems in both experimental conditions (Cue, No-Cue). Furthermore the provision of a cue listing the relevant hypotheses sharply reduced both the proportion of problems in which an irrelevant hypothesis was observed and the proportion of irrelevant hypotheses sampled across problems in both age groups. Nevertheless, neither age group completely excluded irrelevant information from hypothesis sampling even when a cue card was provided.

Hypothesis Sampling Systems: Of major interest to this study was the manifestation, as a function of experimental condition, of the patterns of responses identified as hypothesis sampling systems (Gholson, et al., 1972). The use of verbal probes limited the possible response sequences that could be observed to those representing the three strategies (focusing, dimension checking and hypothesis checking), one of the stereotypes (stimulus preference), and the hypothesis sequences identified as unsystematic (random).

Figure 6 presents age comparisons of the relative frequencies of the various systems for each experimental condition. The first set of five columns, representing the distribution when only relevant stimuli were presented without a cue (Four-Dimension-No-Cue), shows that 45% of third-grade response sequences were categorized as dimension checking. Focusing and the unsystematic category were next in frequency with approximately 20% of the response sequences in each category. Hypothesis checking and stimulus preference respectively, accounted for 8% and 5% of the sequences. In contrast, among the eighth-grade children, the response sequences were almost equally divided between focusing (41%) and dimension checking (45%) with the remaining sequences falling equally

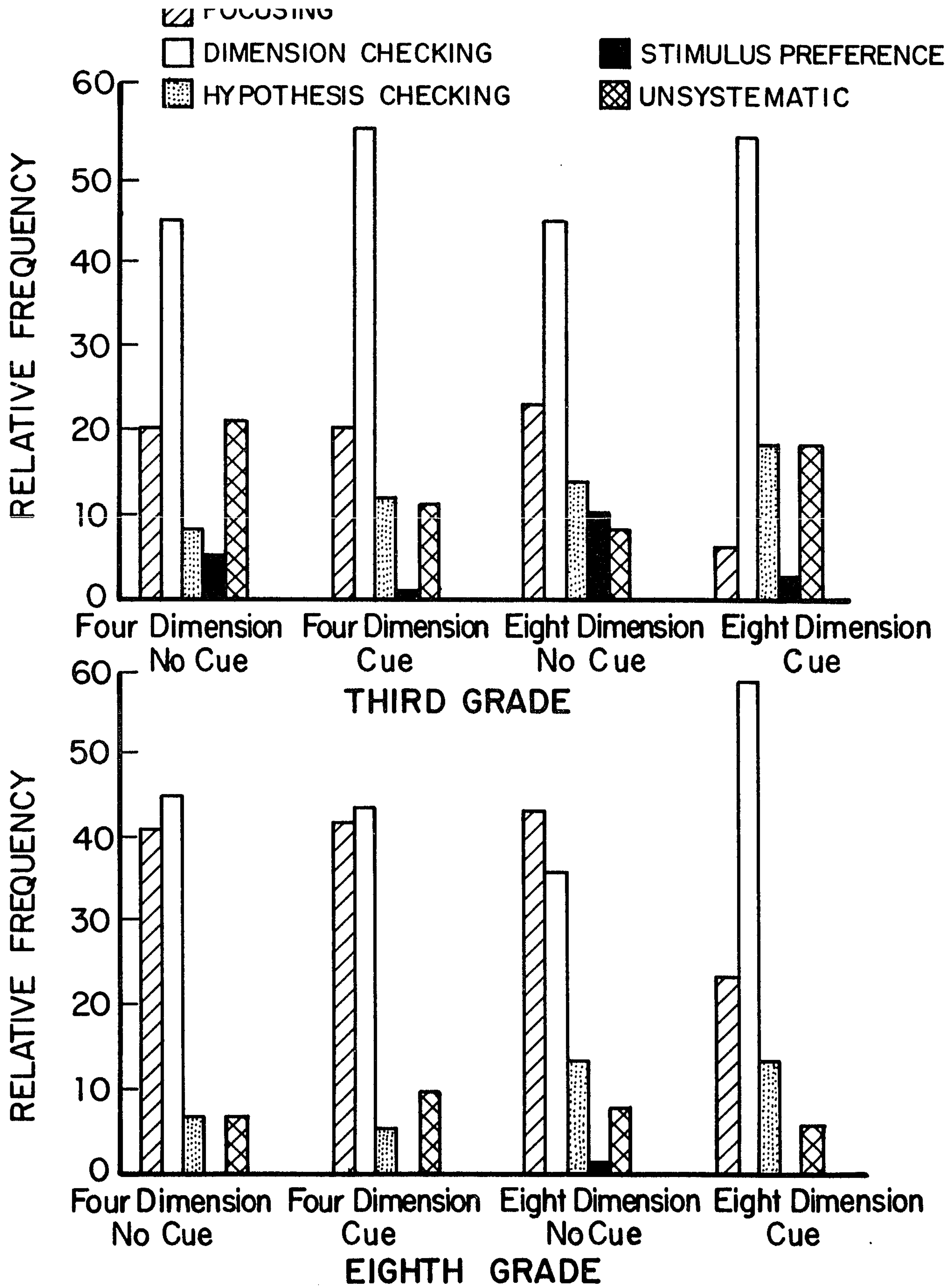


Fig. 6 The frequency distribution of hypothesis sampling systems for each experimental group

into the hypothesis checking and unsystematic categories. The second group of columns on that figure relate to the Four-Dimension-Cue condition. The provision of a cue list with this stimulus configuration did not effect the proportion of third-grade focusing sequences. However, the proportion of dimension checking sequences was increased to 56%. There was also a slight increase in hypothesis checking (4%) and a slight decrease in stimulus preference (4%) but the major decrease was in unsystematic patterns (10%). The cue had virtually no effect on the eighth-grade distribution.

The third set of columns on Figure 8 refers to the distribution of problems in the Eight-Dimension-No-Cue condition in which no irrelevant hypothesis was manifested in the first three trials. The distribution of focusing and dimension checking sequences in the third-grade group was very similar to that of the third-grade Four-Dimension-No-Cue group. The only changes occurred in the other three categories where there was a 6% increase in both hypothesis checking and stimulus preference and a 13% decrease in the unsystematic category. In this condition, the eighth-grade children produced a smaller proportion of dimension checking sequences (36%) than the comparable Four-Dimension-No-Cue group and there was a slight increase in the proportion of all of the other sequences.

Providing a cue with eight dimensional stimuli increased the proportion of dimension checking in both age groups (third-grade; 55% vs 45%, eight-grade; 59% vs 36%) and decreased the proportion of focusing (third-grade; 6% vs 23%, eight-grade; 23% vs 43%). In the third-,but not eighth-grade group, there were also differences in the proportions in the other three categories. Hypothesis checking and unsystematic responding increased (4% and 10% respectively) while stimulus preference decreased (8%) (See Appendix G for all calculations).

These distributions suggest that the presence of irrelevant information in the stimulus configuration without the provision of a cue list, had little effect on the production of strategies in either of the two age groups. But such conclusions are unwarranted as the proportions in Figure 6 are completely interdependent and based on the number of problems that were classifiable. There were dramatic differences in the number of problems in each group that could be classified. As Table 7 reveals, a larger proportion of eighth-grade than third-grade problems in each condition were classifiable. The largest proportion of eighth-grade classifiable problems were found in the Four-Dimension-No-Cue condition (86%). Providing a cue with that configuration led to a small decrease in classifiable problems (82%). Conversely, among the third-grade groups, there was an increase in the proportion of classifiable problems with the provision of a cue (74% vs 80%). Thus, in the Four-Dimension-Cue condition, the distribution includes approximately the same proportion of problems for both age groups. When irrelevant information was included in the stimulus configuration, there was a sharp drop in the number of problems that could be classified in both age groups (third-grade; 40%, eighth-grade; 61%). The proportions were higher when a cue was provided but they remained smaller than any of the Four-Dimension groups (third-grade; 58%, eighth-grade; 64%). Therefore, it is possible that the frequency distributions used in the aforementioned group comparisons do not accurately depict the effect of including irrelevant information in the stimulus complex on the problem solving behavior of children in these two age groups.

Table 7

Proportion of Classifiable Problems for Each Experimental Group

Condition	Grade			
	Third		Eighth	
	No-Irrelevant	Irrelevant	No Irrelevant	Irrelevant
Four-Dimension-No Cue	.737	-	.859	
Four-Dimension-Cue	.801	-	.818	
Eight-Dimension-No-Cue	.398	.702	.609	.807
Eight-Dimension-Cue	.578	.745	.641	.708

Statistical analyses of focusing and dimension checking: It is important to note that the classification of a problem as either focusing or dimension checking was based on a statistical adjustment applied to the actual number of these patterns observed (Levine, 1975, Appendix A). With the feedback patterns used in this study, an apparent focusing sequence could result from a dimension checking plan [1]. Therefore an adjustment was made on the number of focusing and dimension checking sequences observed in each experimental group, based on the number of problems showing a dimension checking sequence in that group. This results in a decrease in the proportion of problems classified as focusing and a concomitant increase in the proportion of problems classified as dimension checking see Appendix H for adjustment formulae and rationale). Focusing and dimension checking accounted for the majority of response sequences in both age groups in all conditions. Therefore, it was of interest to statistically evaluate the similarities and differences in the frequency of each of these systems among the eight groups.

Since the proportion of each of these systems used in the frequency distributions was based on adjustments made on the number of such response sequences observed in each group, rather than the number of such sequences manifested by each subject, it was impossible to obtain a measure of the number of problems per child that fell into each category. However, there is a procedure which leads to a valid evaluation of the statistical significance of differences in the manifestation of each of the systems observed among the various groups. Using this procedure,

1. In the -+- feedback pattern problems, adjustments include one hypothesis checking sequence.

the 24 subjects in each experimental group were randomly divided into subgroups of eight. The statistical adjustments on focusing and dimension checking were performed on each subgroup separately. Following statistical adjustment, the number of problems per subgroup categorized as involving each system was treated as one observation. Analyses of variance (2 x 2 x 2) were based on the three observations per experimental group. These analyses were performed on the number of problems showing focusing sequences and the number of problems showing dimension checking sequences. For focusing sequences, both age (F (1, 16=34.6) and information (F= 22.9) were significant. There were no other significant effects. For dimension checking, both information (F=6.6) and cue (F=12.9) were significant and the information x cue interaction approached significance (F=3.8, p.10< p< .05). The Tukey b procedure, which is particularly conservative with these data as there are only 24 observations across all groups, was applied. With this method, the difference between the two age groups in the manifestation of focusing sequences was significant in the Four-Dimension-No-Cue condition as were the differences between the eighth-grade Four-Dimension groups (No-Cue and Cue) and the eighth-grade Eight-Dimension-Cue group. None of the group comparisons of dimension checking sequendes reached statistical significance with this procedure.

Figure 7 graphically presents the mean number of focusing and dimension checking sequences per experimental condition used in these analyses. A comparison of third- and eighth-grade data reveals that the older children manifested a larger number of focusing sequences than the younger children in each condition. In both age groups, the number was reduced when irrelevant information was included in the stimulus complex and further reduced when a cue was provided with that configuration.

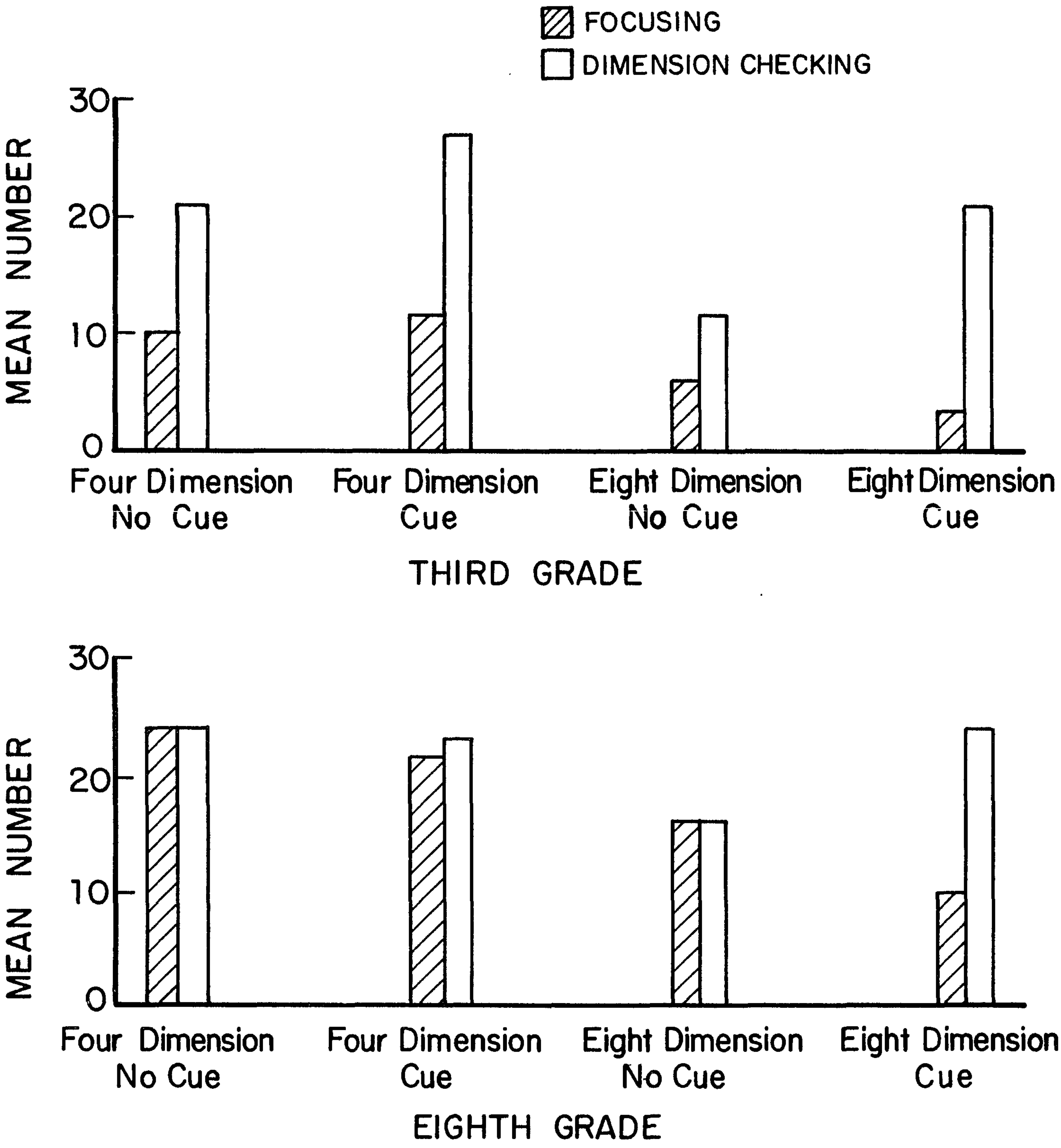


Fig. 7 Mean number of focusing and dimension checking for each experimental group

Note: Based on three randomized observations.

The number of dimension checking sequences in both age groups was also reduced when irrelevant information was included in the configuration. However, in contrast to focusing, the provision of a cue with both the relevant and irrelevant stimulus complex resulted in an increase in dimension checking in both age groups. The near significant information x cue interaction on the dimension checking measure appears to be a function of the differential effect of the cue on the manifestation of this strategy by the two age groups when irrelevant information was included in the stimulus complex versus when it was not. Whereas the number of dimension checking sequences increased in both age groups when a cue was provided with the relevant-irrelevant stimulus complex, a similar effect was observed in only the third-grade group when the stimuli consisted of only relevant dimensions.

Thus, when the number of problems classified in each of these system categories is used as basic data rather than the proportions dependent on the number of classifiable problems which are used in the systems distribution, a different picture of the effects of including irrelevant information in the stimulus complex emerges. In both age groups, when irrelevant information was included in the configuration, there was a reduction in the number of problems showing focusing and dimension checking sequences. However, the relative occurrence of each remained fairly stable across both Four-Dimension conditions (No-Cue and Cue) and the Eight-Dimension-No-Cue condition; i.e., among third-grade groups, the mean number of dimension checking sequences was approximately double that of focusing sequences in all three conditions whereas among eighth-grade groups, the mean number of dimension checking and focusing sequences were approximately equal in these same conditions. Providing a cue when irrelevant information was included in the configuration resulted in a decrease in focusing and an increase in dimension checking in both age groups. Thus, in the Eight-Dimension-Cue condition the mean number of problems among the third-grade group that showed a dimension checking sequence was approximately seven times that of the number of problems that showed a focusing sequence. The only eighth-grade group in which the mean number of problems showing focusing and dimension checking sequences were not approximately equal was the Eight-Dimension-Cue group. Here the mean number of problems showing a dimension checking sequence was more than double the number showing a focusing sequence.

The distribution of Hypothesis Sampling Systems when problems containing irrelevant hypotheses are included: In the eight dimension conditions, a number of problems (see Table 7) showed the same kinds of sequences of hypotheses as those included in the previous computations, except that one or more of the first three hypotheses were values from dimensions designated as irrelevant. If the pattern of hypotheses follows the dimension checking, hypothesis checking, stimulus preference or unsystematic pattern, it seems reasonable to assume that the same problem solving process underlies each specific hypothesis sequence regardless of whether irrelevant information is included or excluded from sampling. Therefore a new distribution was calculated for each of the Eight Dimension conditions. Those problems which showed irrelevant hypotheses as one or more of the hypotheses following the first three feedback trials in one of the aforementioned response patterns were added to the appropriate category. It is, however, questionable whether a problem in which irrelevant hypotheses following feedback trials 1, 2 or 3 are manifested can be assumed to indicate a focusing plan even if each of the hypotheses are globally consistent and the solution hypothesis is verbalized following the third feedback trial. Both the size of the set from which the subject initially samples and the elimination process following feedback are unknown. Therefore, including problems that show such a sequence in the focusing category may be in error. Therefore, those globally consistent sequences that showed the solution hypothesis following the third feedback trial but also included irrelevant hypotheses were added to the dimension checking category.

Comparison of the hypothesis sampling distributions when problems with irrelevant hypotheses are included and excluded are presented in Figure 8. Including problems in which irrelevant values were verbalized as hypotheses following the first three feedback trials primarily affected the distribution in the Eight-Dimension-No-Cue condition. In both age groups, the predominant change was an increase in the proportion of problems categorized as dimension checking (third grade; 45% vs 61%, eighth grade; 36% vs 50%) and a decrease in the proportion of problems categorized as focusing (third grade; 23% vs 13%, eighth grade; 43% vs 33%). It might be thought that the decrease in focusing is an artifact of categorizing all globally consistent sequences that included irrelevant hypotheses as dimension checking. However, this procedure only affected a small number of problems. In the Eight-Dimension-No-Cue conditions, there were three of these sequences among all third-grade problems and five among all eighth-grade problems. In the Eight-Dimension-Cue condition, there was one in each age group. In contrast, a dimension checking pattern accounted for the majority of problems that included irrelevant hypotheses in the H_1, H_2, H_3 sequence. As Table 8 reveals, when irrelevant values were sampled in the initial sequence, this pattern was observed on 76% and 75% of the third-grade problems

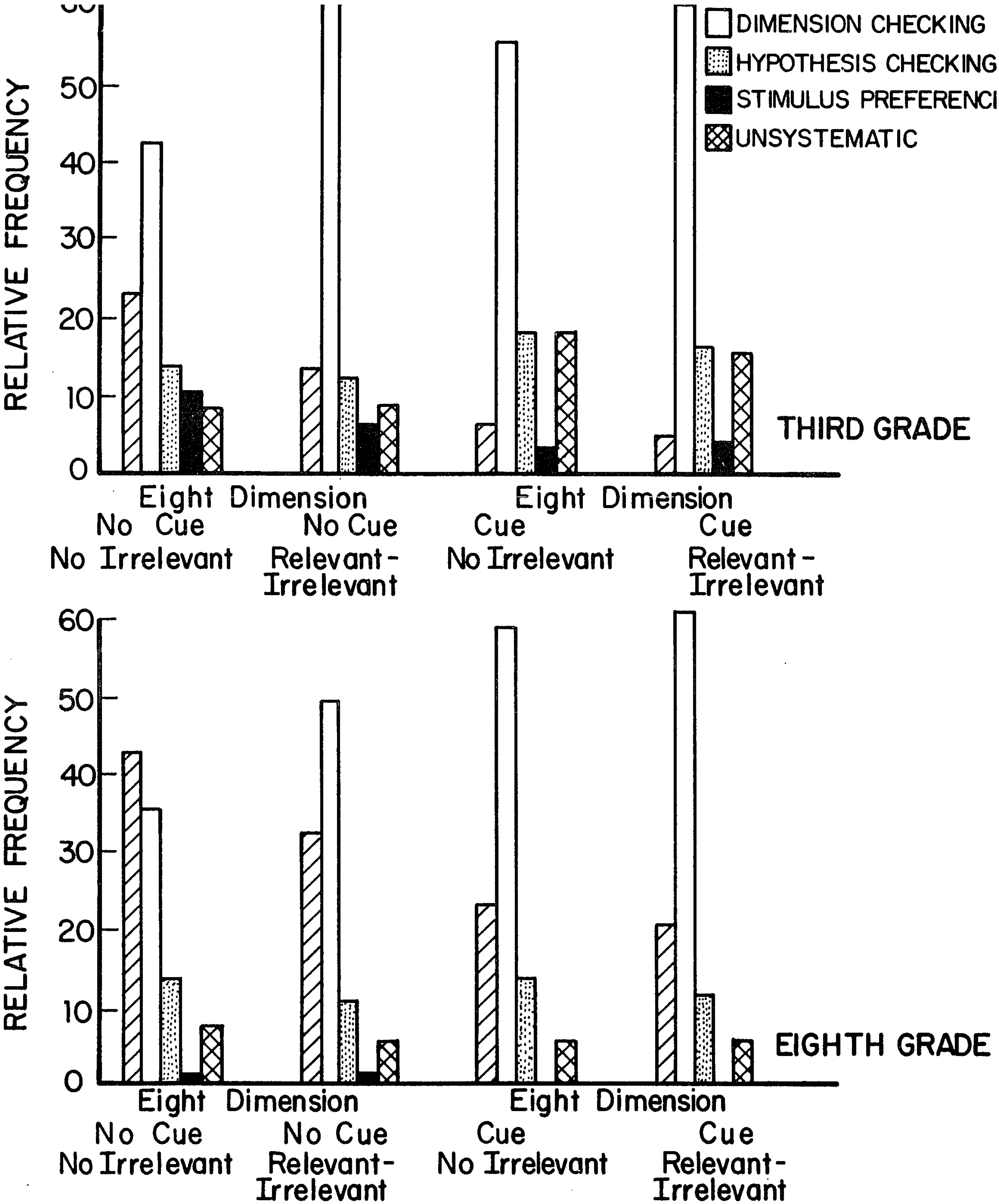


Fig. 8 Comparison of hypothesis sampling systems in eight dimension conditions including & excluding problems showing irrelevant hypotheses

Table 8

Percentage of Problems Showing Irrelevant Hypothesis as the First, Second or Third Hypothesis for Each Response Sequence

Condition	Global Consistency	Dimension Checking	Hypothesis Checking	Stimulus Preference	Unsystematic
Third grade					
Eight-Dimension-No-Cue	5	76	10	0	9
Eight-Dimension-Cue	3	75	9	6	6
Eighth grade					
Eight-Dimension-No-Cue	14	81	5	0	0
Eight-Dimension-Cue	8	85	0	0	8

Note: Number of problems per condition

	Third grade	Eighth grade
Eight-Dimension-No-Cue	58	37
Eight-Dimension-Cue	32	13

before statistical adjustment in the Eight-Dimension-No-Cue and Cue conditions, respectively and on 81% and 85% of the respective eighth-grade problems.

Thus, it appears that in the Eight-Dimension-No-Cue condition, third-grade children who excluded irrelevant hypotheses from sampling tended to follow the same patterns they had when no irrelevant hypotheses were included in the configuration (Four-Dimension-No-Cue). Under the same circumstances, eighth-grade children showed no difference in the manifestation of focusing plans but there was a slight decrease in the manifestation of dimension checking plans. However, when irrelevant hypotheses were included in the sampling set, children in both groups tended to follow a dimension checking pattern. Thus, for both age groups, the total distribution of categorizable problems included a larger proportion of dimension checking plans and a smaller proportion of focusing plans in the Eight-Dimension-No-Cue condition than in either of the comparable Four-Dimension groups.

In the Eight-Dimension-Cue groups, including those problems which showed irrelevant hypotheses sampled in the initial sequence did not appreciably change the distribution. First of all, in both age groups, there were fewer problems in which irrelevant hypotheses were sampled following the first three feedback trials. Secondly, dimension checking was the modal plan observed in the problems of children in both age groups even when no irrelevant hypotheses were sampled following the first three feedback trials. The third-grade groups manifested dimension checking plans in the same proportion of problems as their peers in the Four-Dimension-Cue group and the eighth-grade children showed a higher proportion of dimension checking plans in the Eight-Dimension-Cue group than in any other group. Therefore, including problems in which irrelevant hypotheses were sampled in the Eight-Dimension-Cue distributions served mainly to accentuate the relative frequencies observed when these problems were excluded from the distribution.

Categorization of all problems: As Table 7 indicates, there were still a number of problems that did not meet criteria for classification in any hypothesis sampling system category (Gholson et al., 1972). An attempt was made to modify the systems categories so that all problems could be included in the frequency distribution. It was thought that this inclusion might bring the age and experimental differences into sharper focus.

The Gholson et al. (1972) systems analysis excludes from consideration all problems in which an hypothesis followed by positive feedback on Trial 2 is not maintained. Earlier it was shown that such shifts may not always represent insensitivity to feedback. In fact, these shifts may indicate that the child is attempting to focus: i.e., trial 1, the positive stimulus is 'black, big, square, L', the first hypothesis is 'black'; trial 2, the positive stimulus is 'black, small, square, T'. At this point, both 'black' and 'square' have appeared in the two positive stimuli and either could be the solution. If the subject shifts from 'black' to 'square', receives negative feedback following his choice response and chooses 'black' he will have reached solution as 'black' and 'square' appear on different stimuli on the third trial. Problems showing this pattern were added to the focusing category after the statistical adjustment made on the observed regular focusing and dimension checking patterns.

On the other hand, a number of problems are included in the hypothesis sampling system adjustments as focusing which show a focusing sequence but which actually appear to be overt examples of 'lucky' dimension checking. In these problems, the third hypothesis is the logical solution but even though positive feedback follows the choice a different hypothesis is verbalized following feedback. Since the underlying process assumed to be operating in focusing is a plan of sequential elimination which leaves the subject with only the logical solution as the third hypothesis, shifting from solution suggests the subject was not operating with that plan. Thus, in these data, all problems which showed focusing sequences in which the solution was not held though criterion were added to the dimension checking category. Thus, the dimension checking category includes those sequences identified by Gholson et al. (1972) for each of the feedback patterns plus those sequences which appear to follow a focusing plan but show a shift from solution subsequent to the third hypothesis. These problems were included in the statistical adjustments on dimension checking and focusing. Another group of problems was excluded from the systems analysis that involve shifts following Trial 2 positive feedback. These appear to follow a dimension checking plan. Unlike the shifts in the problems placed in the focusing category, the shift is not necessarily globally consistent but is always locally consistent. For example, the subject receives positive feedback for his choice of the configuration containing the following cues; 'black', 'circle', 'dashed' and 'up'. He, then verbalizes 'black' as his hypothesis and chooses the configuration 'black', 'square', 'up' and 'solid' on the second feedback trial. He then verbalizes 'square' as his second hypothesis. That hypothesis is locally consistent as it appeared on the preceding positive stimulus but it is not globally consistent as it did not appear on the first positive stimulus. This sequence appears to follow the same pattern as the dimension checking sequences included in the distribution; i.e., a series of locally consistent hypotheses, each from a different dimension. This seems to be a variant method of sampling dimensions and therefore problems showing such sequences were added to the dimension checking category. Also, in the Eight-Dimension conditions, all dimension checking and focusing sequences that included irrelevant hypotheses were added to this category.

The hypothesis checking and stimulus preference categories are comprised of the sequences identified by Gholson et al. (1972) with and without the inclusion of irrelevant hypotheses.

The unsystematic category designated by Gholson et al. (1972) includes three different kinds of unsystematic sequences. One involves a sampling of hypotheses which does not follow any classifiable sequential pattern. However, none of the hypotheses shows an insensitivity to feedback or a loss of memory for hypotheses previously verbalized and disconfirmed during the three hypotheses sequence used for classification (i.e.,no locally inconsistent hypotheses, no repetition of previously disconfirmed hypotheses). Thus, if the first trial response receives positive feedback and negative feedback is given on the second and third trials, the sequence is 'black', 'big', 'small' or 'black', 'big', 'white'. In the following distributions, sequences of this kind comprise the unsystematic category. However, two other kinds of sequences included in Gholson's (1972) unsystematic category involve hypothesis testing 'errors'. In these sequences, there is either insensitivity to feedback or loss of memory for hypotheses verbalized and disconfirmed on the preceding trials. In these cases, if the first trial response receives positive feedback and negative feedback is given on the second and third trial, the patterns might be variations of 'black', 'black', 'big' or 'black', 'big', 'black'. Sequences such as these do not, in principle, appear to differ from a number of other sequences excluded from systems analysis because they include locally inconsistent hypotheses. Therefore, problems with unclassifiable sequences which include the aforementioned hypothesis testing 'errors' were classified as unsystematic error.

Within the group of problems excluded from the systems analysis are a number of problems which include a locally inconsistent hypothesis but appear to be systematic. It was noted earlier that one of the errors that led to non-solution was testing a locally inconsistent hypothesis drawn from the solution dimension. In this study, there were a sizable number of problems in which an apparent dimension checking sequence was observed but one or more locally inconsistent hypotheses were manifested. In this case, the sequence may reflect sensitivity to feedback but not to stimulus information. It appears that the subject is sampling hypotheses dimensionally but is inattentive to or forgets which hypotheses appeared on the last positive stimulus (i.e., were locally consistent). In this analysis, these errors are classified as dimension checking errors.

Thus, in Figure 9, all problems are categorized as either focusing, dimension checking, hypothesis checking, stimulus preference, unsystematic, dimension checking error, or unsystematic error. The frequency distribution of strategies when all problems are included, mirrors the distribution calculated by the Gholson et al. (1972) system with, of course, smaller relative frequencies in each category. (See Appendix G for all calculations).

The modal strategy of third-grade children in each condition was dimension checking. In the various conditions, this system accounted for 37% to 47% of the problems. Only a small proportion of problems in any condition showed a focusing approach. The range was from 3% (Eight-Dimension-Cue) to 18% (Four-Dimension-Cue). The proportion of problems showing hypothesis checking sequences ranged from 6% (Four-Dimension-No-Cue) to 12% (Eight-Dimension-Cue).

Taken together, strategic approaches to problem solution among third-grade children ranged from 58% (Four-Dimension-No-Cue) to 74% (Four-Dimension-Cue) among third-grade children. Thus, the problem solving approach of children in this age group was most efficient when the stimulus contained no irrelevant information and a cue list was provided. The primary effect of including irrelevant information in the stimulus was a reduction in the proportion of problems showing a focusing approach. But, even in the Four-Dimension-Cue group, this approach was only observed on 18% of the problems.

The eighth-grade children, on the other hand, showed the most efficiency with four dimensional stimuli when no cue was provided. Providing a cue did not change the bimodal distribution (approximately equal proportions of focusing and dimension checking) but it did decrease the total strategic approaches from 84% to 78%. Both eighth-grade groups exposed to configurations that contained irrelevant stimulus information showed strategic approaches on 76% of the problems. In the Eight-Dimension-No-Cue group, the primary decrease was in focusing plans which was observed on 27% of the problems in contrast to 38% of the problems in the Four-Dimension-No-Cue group. When the cue was provided with a configuration

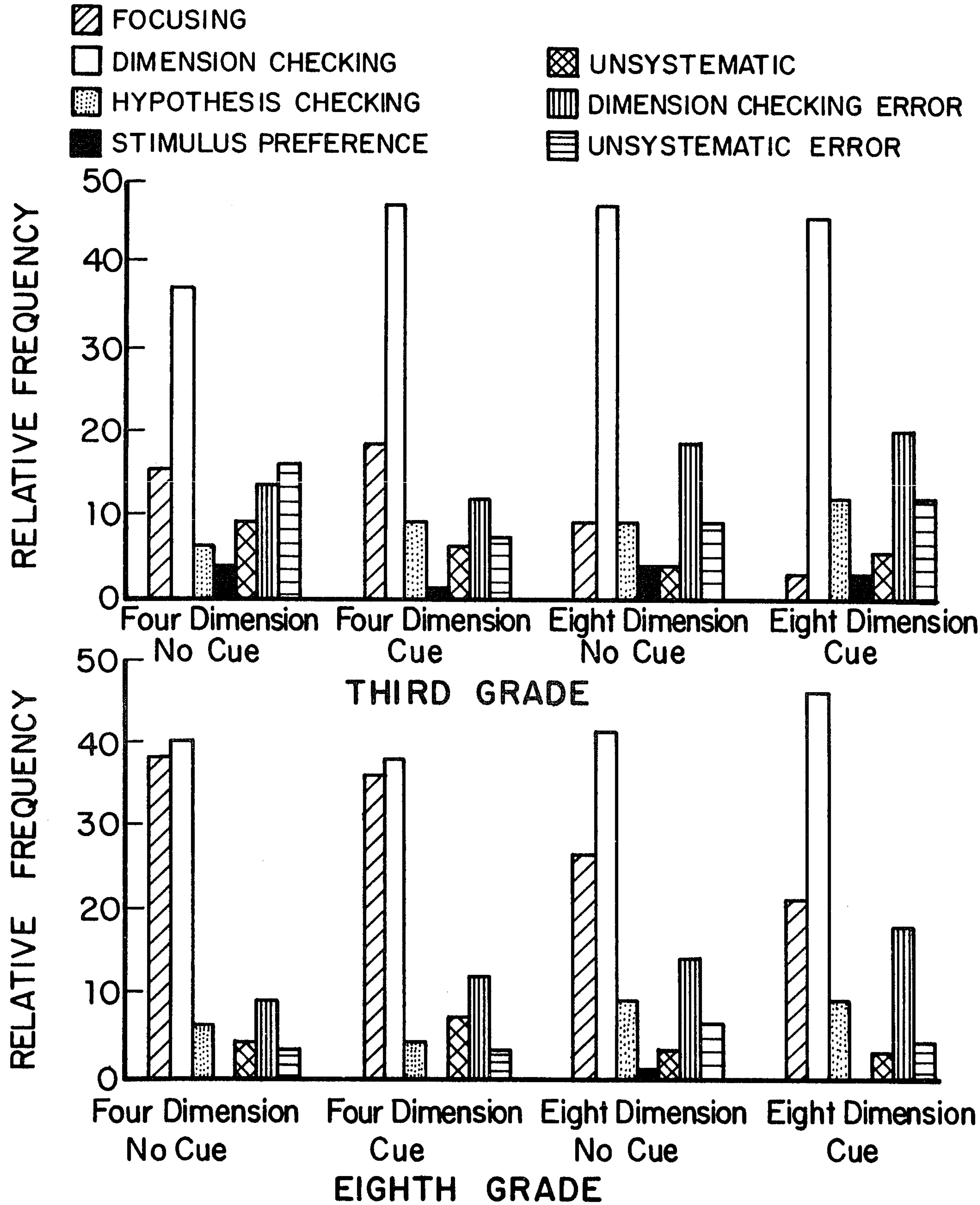


Fig. 9 The frequency distribution for each experimental condition when all problems are categorized

containing irrelevant cues, the proportion of focusing was further reduced to 21%. At the same time, there was a concomitant increase in dimension checking (46% versus 38-40%) in the other three conditions.

In those problems in which a strategic approach was not evidenced, error patterns, rather than stereotypic or unsystematic patterns, were likely to be observed. Among the third-grade groups, error patterns were observed in from 19% (Four-Dimension-Cue) to 32% (Eight-Dimension-Cue) of the problems. Unsystematic error patterns were most evident in the Four-Dimension-No-Cue condition (16%) and least evident in the Four-Dimension-Cue and Eight-Dimension-No-Cue condition (7% and 9%). Dimension checking errors were most evident when irrelevant information was present (Eight-Dimension-No-Cue 18%, Eight-Dimension-Cue 20%) and there was little difference in the manifestation of this pattern in the Four-Dimension conditions (No cue 13%, Cue-12%).

The eighth-grade children showed few unsystematic error patterns across conditions (3% to 6%). Dimension checking errors were least evident in the Four-Dimension-No-Cue condition (9%). There was a slight increase with the provision of cue (12%). Including irrelevant information in the configuration without a cue also resulted in an increase (14% and the highest proportion was found in the group provided with a cue with a relevant-irrelevant stimulus complex (18%).

Comparison of focusing and dimension checking problem solving processes. It is possible to question whether focusing and dimension checking involve completely different problem solving processes. It is assumed (see Appendix A) that in a focusing approach, the subject utilizes a plan in which he initially encodes the total set of hypotheses then sequentially eliminates subsets of these that are logically disconfirmed following each feedback trial. In contrast with dimension checking it is assumed that the subject initially notes the potential hypotheses, codes them in dimensional sets of two, and manifests only one of them. Following each negative feedback trial, he eliminates a dimension. When he chooses a new dimension to test, his hypothesis is the cue on the dimension consistent with feedback given on that trial (i.e., is locally consistent). He is assumed to either keep track of the dimensions previously tested and disconfirmed or of those that remain to be tested. But a subject would appear to manifest dimension checking if he encoded the four hypotheses on the first positive stimulus and tested each of these sequentially until solution was achieved. This would resemble a focusing approach, but it involves a simpler comparative process. Unlike the mental process of logical intersection in which <u>subsets</u> of hypotheses are eliminated following <u>each</u> feedback trial as assumed in a focusing plan, this plan would involve manifestation of each hypothesis which appeared in the first positive stimulus. Rejection of a dimension and choice of a new hypothesis from a dimension that has not yet been sampled would take place, of course, following only <u>negative</u> feedback trials. If dimension checking did involve a long range plan of this type in which first trial information was retained, second and third hypotheses in problems categorized as involving dimension checking would always be cues included in the stimulus that was positive on the first feedback trial. In fact, the contrary was the case: In problems categorized as dimension checking in the systems analysis approximately 65% of all second and third hypotheses were inconsistent with feedback information on the first trial (See Appendix I for figures).

Related evidence is found by examination of the relationship between the third hypothesis and feedback information from the second trial. Here, in approximately 45% of problems showing a dimension checking approach, the third hypothesis was inconsistent with second trial feedback information. Thus, it does not appear that dimension checking involves much retention of specific stimulus information from previous feedback trials. It is clearly not the case that retention of first trial information is an integral part of the dimension checking system. The evidence indicates then, that a dimension checking approach does not involve a plan similar to focusing.

Dimension checking solution: A dimension checking strategy is classified on the basis of hypotheses manifested during the first three trials. At that point, either two or three dimensions (depending on feedback condition) have been overtly sampled. It is of interest to ascertain if those problems that are begun with a dimension checking plan are solved without deviation from that plan. That is, the subject solves the problem manifesting only one hypothesis per dimension, without repeating previously disconfirmed hypotheses and without testing locally inconsistent hypotheses. Figure 10 presents the proportion of problems classified unambiguously as involving dimension checking that were solved dimensionally without any of the aforementioned errors in each experimental condition.

It may be seen that the eighth-grade children solved dimensionally more of the four dimension problems begun with this plan that did the third-grade children. Although the eighth-grade children were no more likely to use a dimension checking plan when a cue was provided with this configuration than when it was not (See Note, Figure 10), they were somewhat more likely to successfully carry this plan to conclusion without error when the cue card was provided (59% vs 69%). While the third-grade children were more likely to use a dimension checking plan when a cue was provided, there was little increase in the likelihood of these children reaching solution without error with that plan (40% to 43%).

A larger number of dimension checking plans were observed among both age groups when irrelevant information was included in the stimulus complex than when it was not (see note, Figure 10). However, among both age groups, a smaller proportion of these were solved dimensionally and there was no difference as a function of age (24%). Although providing a cue list with this configuration resulted in a larger proportion of successful completions in both age groups, there was no difference as a function of age (46%). However, whereas among third-grade groups this represented a slightly higher ratio of dimension checking solution to dimension checking approach than had been evidenced in the Four-Dimension conditions, the proportion observed in this eighth-grade group remained below that of the eighth-grade Four-Dimension groups.

A note of warning should be attached to interpretation based on the preceding data. The baseline reflects only problems in which a dimension checking approach was observed. Recall that this only represents a portion of those problems categorized as dimension checking. Theoretically, it is assumed that a specified (according to feedback pattern) proportion of globally consistent sequences result from 'lucky' dimension checking plans but the solution hypothesis is drawn from the second or third dimension sampled. Therefore the baseline used in these calculations refers to a smaller proportion of problems in each group than were actually categorized as dimension checking. Moreover, the proportion of

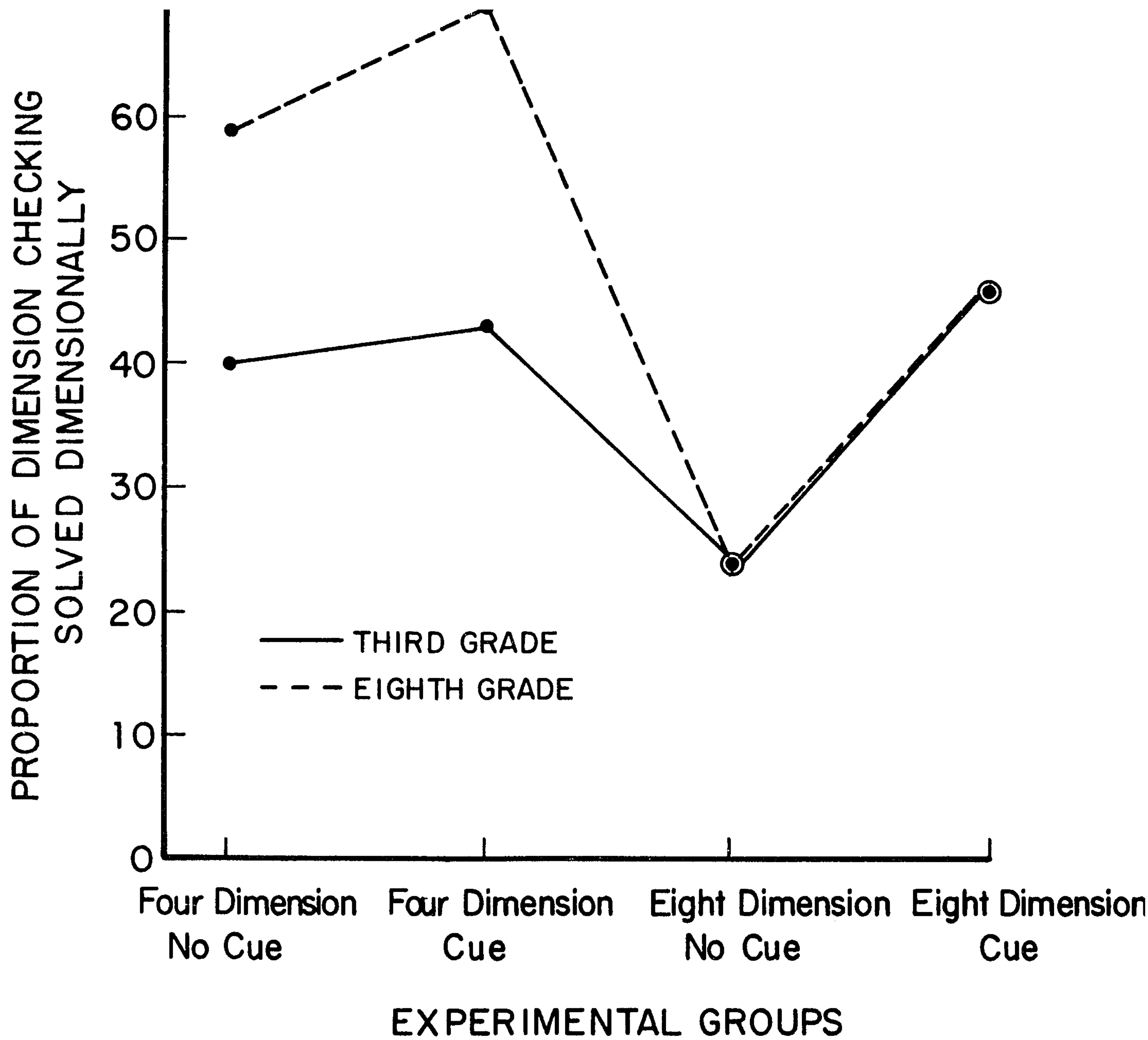


Fig. 10 Proportion of problems classified unambiguously as dimension checking solved dimensionally by each experimental group

Note: Number of problems classified unambiguously as dimension checking for each experimental group.

	THIRD GRADE	EIGHTH GRADE
Four Dimension No Cue	47	49
Four Dimension Cue	58	48
Eight Dimension No Cue	72	58
Eight Dimension Cue	68	61

problems used as the baseline differed among subjects and no statistical analysis was appropriate. Thus, the preceding analysis must be considered a rough but conservative estimate of the proportion of problems solved without error using the dimension checking plan.

<u>Effects of feedback pattern</u>: Levine (1966) has shown that among adults global consistency is an increasing function of the number of positive feedbacks prior to negative feedback on the third trial. In this study, two problems presented each child yielded two such positive feedback trials (++-), two problems yielded two negative feedback trials (---) and four of each subject's problems consisted of one negative and one positive feedback trial preceding negative feedback on the third trial (+--, -+-). Since preliminary inspection of the data suggested that the effect observed in Levine's adults was also present in the protocols of subjects in these two age groups the proportion of globally consistent sequences under each feedback pattern was calculated for each experimental group. As Figure 11 indicates, among children of both ages, the largest proportion of globally consistent hypothesis sequences was observed in response to the ++- feedback pattern (eighth - 55%, third 34%) and the lowest in response to the --- feedback pattern (eighth - 29%, third - 16%). Further, the mixed feedback patterns (-+-, +--) in each case were intermediate. (See Appendix J for the percentages per experimental group).

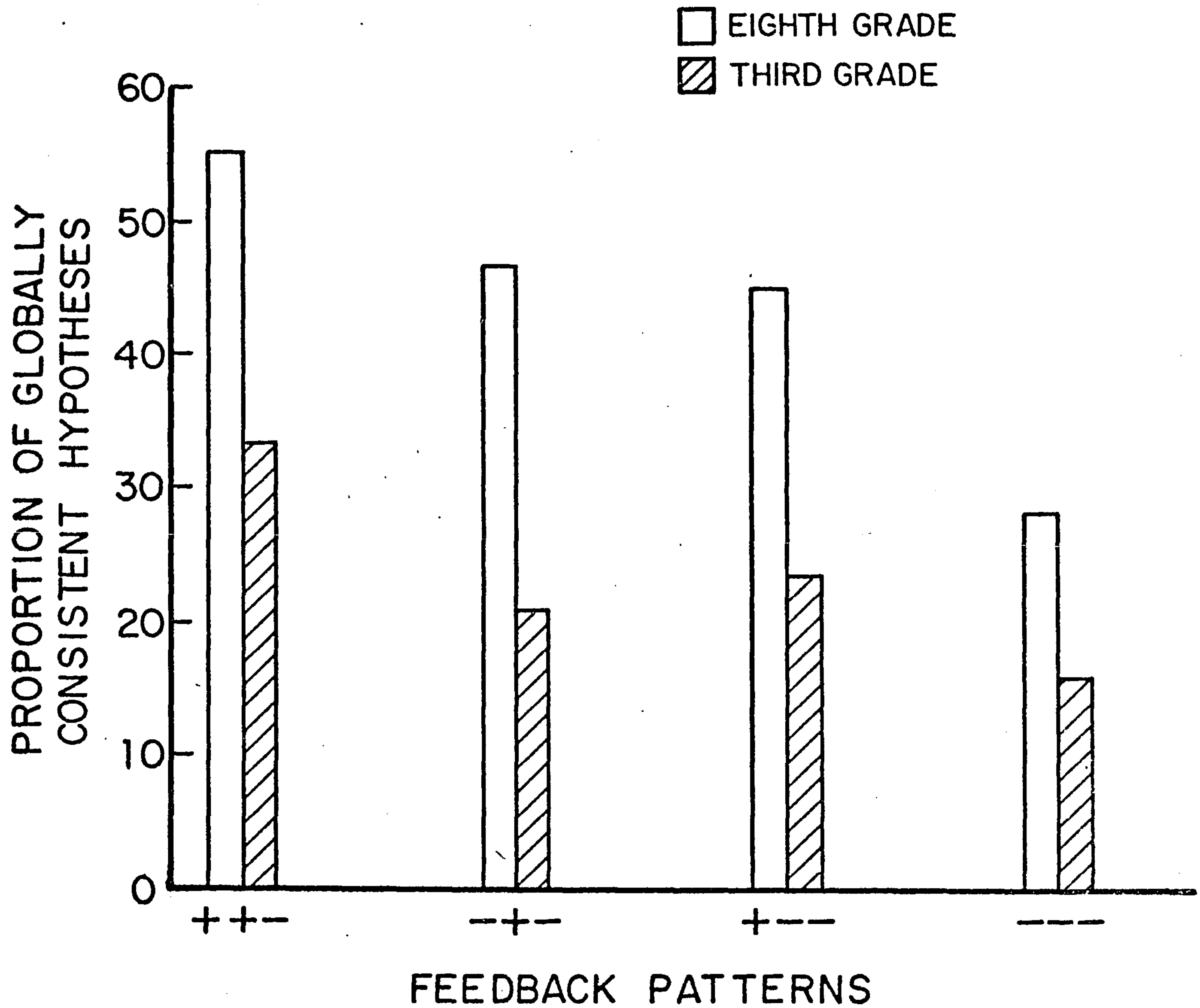


Fig. 11 The proportion of globally consistent hypotheses sequences across all experimental conditions for each feedback pattern

Recall that in this study, a statistical adjustment was made on those problems which showed a globally consistent hypothesis sequence in order to eliminate the possibility of misinterpreting 'lucky' dimension checking plans as exemplars of a focusing approach. The adjustment was made separately for each feedback pattern (Levine, 1975, Appendix A). Thus it was possible to calculate for each of the three randomized subgroups in each experimental condition the proportion of focusing problems per feedback pattern. Using these figures, a series of 2 x 2 x 2 (Age x Information x Cue) analyses of variance were performed on the differences between each feedback pattern and the mean of the other three groups (e.g., ++-/means ---, -+-, +-- and ++-/---). None of the results of the five possible comparisons was significant. However, as Table 9 indicates, standard tests for the grand mean over all conditions and both ages (Mood, 1950) revealed statistically significant differences between ++- and ---, ++- and means for the other three feedback patterns and --- and means for the other three feedback patterns. Thus, significantly more focusing patterns were observed in response to ++- feedback than to any other feedback and significantly fewer focusing patterns were observed in response to --- feedback than to any other feedback pattern. This indicates that the use of a focusing strategy is, at least, partially a function of the particular feedback pattern. It appears that Levine's conclusion in regard to adult problem solving behavior also holds for problem solving behavior of third- and eighth-grade children.

Table 9

Standard Test for the Grand Mean Results for Differences in Focusing as a Function of Feedback Pattern

Differences	F
++- $\bar{x}$ -+-, +--, ---	10.11*
--- $\bar{x}$ ++-, -+-, +--	12.79*
+-- $\bar{x}$ ++-, -+-, ---	.70
-+- $\bar{x}$ ++-, +--, ---	.26
--- ++-	18.89*

*$P < .05$

Verbalization of Problem Solving Techniques

Following the last problem, each child was asked "How did you figure out the answers to these problems?" The responses were coded first into 36 categories based on the procedures the children stated they had followed. These categories were then reduced to 14 in relation to the underlying theme expressed (e.g., elimination of hypotheses, memorization of specific stimulus information, attention to feedback information, etc.).

It is assumed that these verbal reports are only a gross indication of the problem solving techniques used by the children: that they represent only the most salient aspect of the procedure, rather than the actual or complete procedure followed, and that language facility differed between the two age groups. Thus, the verbalizations of the older children were expected to be more explicit. (The frequency count for each category and an exemplar of each may be found in Appendix K.)

The following is a summary of the most striking thematic differences verbalized in the two age groups. 45 eighth-grade children and 2 third-grade children expressed some variation of the notion of eliminating hypotheses; 30 of these 45 eighth-grade children but neither of the 2 third-grade children stated that elimination was based on the nature of the feedback sequence on each trial beginning with the first. Eighth-grade children also more frequently expressed the idea that memory was a factor. Some mention of memory was made by 33 third-grade children and 69 eighth-grade children. Of these, 12 third-grade children and 43 eighth-grade children specifically mentioned attempting to retain first trial positive stimulus information. Nine eighth-grade children and no third-grade children specifically mentioned checking dimensions.

On the other hand, third-grade children outnumbered eighth-grade children in stating that their attention was directed to positive stimulus information (20 third-graders, 10 eighth-graders). Third-grade children were also more likely to indicate that their technique consisted solely of changing hypotheses based on feedback. Explanations which involved sampling hypotheses only, were given by 10 third-grade children and two eighth-grade children. Fourteen third-grade children, but no eighth-graders, failed to provide any explanation.

In those conditions in which a cue was provided, 15 third-grade children and 7 eighth-grade children mentioned using the cue list. Of these, 7 third-grade and 3 eighth-grade children verbalized this as the only technique used. In the Eight-Dimension-No-Cue condition, 13 third-grade children and 7 eighth-grade children mentioned attempting to remember the relevant set of hypotheses. Of these, 9 third-grade children but no eighth-grade children gave this as the sole explanation of their procedure.

The previously mentioned unreliability of post experimental reports limit the significance that may be attached to these age differences. Nevertheless, it is interesting that eighth-grade children were more likely to verbalize sequential processing of information whereas third-grade children were more likely to stress attention to individual pieces of information (i.e.,specific feedback and type of stimulus information.)

Summary and Concluding Discussion:

In all experimental conditions, children of both age groups were more likely to use efficient hypotheses sampling techniques than not use them. When a choice response received positive feedback, an hypothesis was retained 93% to 98% of the time. When a choice response received negative feedback, the previous hypothesis was retained only 1% to 5% of the time. Following negative feedback, hypotheses were consistent with feedback (i.e., locally consistent) 79% to 93% of the time. When a child changed an hypothesis, the hypothesis verbalized was one which had previously been sampled and disconfirmed only 6% to 21% of the time. When irrelevant hypotheses were available for sampling, they were sampled during 11% to 59% of the problems and 4% to 28% of the trials.

Children in both age groups were also more likely to show efficient sequential approaches to problem solution than to show unsystematic or error sequences. Across the various conditions, strategic problem solving approaches were evidenced during 59% to 84% of the problems. Dimension checking plans were completed without error on 24% to 69% of the problems showing that plan. Between 53% and 94% of all problems were solved.

Similar types of errors were shown by children in both age groups. Failure to sample the solution dimension was the most common error in unsolved problems. Between 4 to 65 problems per condition showed this type of error. Children in the two age groups also reacted similarly to the various feedback sequences. Problems which were characterized by positive feedback on the first two trials were most likely to show a focusing approach whereas problems initiated by two negative feedbacks were least likely to show this approach.

Despite the similarities in performances among children of the two age groups, there was a significant main effect of age on all measures of hypothesis testing efficiency. Eighth-grade children were more likely than third-grade children to retain an hypothesis following positive feedback and less likely to retain an hypothesis following negative feedback. Following negative feedback, they were more likely to sample a locally consistent hypothesis and when they changed hypotheses, the new hypotheses was less likely to have been previously sampled and disconfirmed. Similarly, when irrelevant information was included in the stimulus complex, the eighth-grade children sampled fewer irrelevant hypotheses during fewer problems.

The older children also showed strategic approaches during a greater proportion of problems than the younger children. They showed fewer unsystematic hypothesis sequences, fewer error sequences and fewer stereotypic sequences. The proportion of problems showing the most efficient strategic approach, focusing, was significantly greater among the eighth-grade children. In addition the eighth-grade children solved significantly more problems than third-grade children.

The difference between the two age groups was most pronounced in the Four-Dimension-No-Cue condition. All age comparisons involving simple hypothesis testing measures yielded significant differences within this condition.

The older children showed a bimodal approach with an equal distribution of focusing and dimension checking, while the third-graders showed little focusing and dimension checking in a little less than half of the problems. Moreover, the older children completed more dimension checking plans without error than the younger. They also solved a significantly larger number of problems.

Providing a memory aid in the form of a cue listing the relevant hypotheses only minimally affected the functioning of eighth-grade children. There was no significant difference in hypothesis testing efficiency or in the proportion of problems solved as a function of the cue. In each condition the distribution was bimodal with the highest proportions equally divided between the focusing and dimension checking categories.

The includion of irrelevant information in the stimulus configuration did interfere with the problem solving efficiency of eighth-grade children. Although they sampled many more relevant hypotheses than irrelevant hypotheses, they did sample irrelevant hypotheses. In comparison to their peers in the Four-Dimension-No-Cue group, eighth-grade children in the Eight-Dimension-No-Cue group showed less efficiency on all measures of simple hypothesis testing efficiency, but only the local consistency comparison reached statistical significance. The children exposed to stimulus configurations containing irrelevant information also solved fewer problems.

In comparison to their peers exposed to only relevant information (eighth-grade Four-Dimension-No-Cue) the eighth-grade group exposed to a stimulus complex containing irrelevant information (Eight-Dimension-No-Cue) showed a strategic approach on fewer problems. The increase in non-strategic approaches was primarily in the dimension checking error category. On those problems which showed no irrelevant hypotheses following the three initial feedback trials, the proportion of focusing sequences was equal to that in the comparable Four-Dimension-No-Cue group but the proportion of dimension checking sequences was lower. The modal sequence observed on those problems in which irrelevant hypotheses were sampled was dimension checking. There were, however, fewer dimension checking plans successfully concluded when irrelevant information was present than when it was not.

The primary purpose of providing a cue list with the relevant-irrelevant stimulus complex was to exclude the necessity of retaining in memory the set of hypotheses relevant to solution of a particular problem. The finding that eighth-grade children showed fewer irrelevant hypotheses on fewer problems when the cue was present suggests that the cue served its purpose. However, some interference in problem solving efficiency is attested to by the finding that the eighth-grade children in the Eight-Dimension-Cue group were less likely to retain an hypothesis following positive feedback than their peers in any other condition. In addition, local consistency did not significantly improve with the provision of a cue list and remained lower than that observed in either of the eighth-grade Four-Dimension groups. However, the eighth-grade children in the Eight-Dimension-Cue condition did solve significantly more problems than those in the Eight-Dimension-No-Cue group, while there was no significant difference between the proportion solved by the Eight-Dimension-Cue group and the two Four-Dimension groups.

The proportion of strategic approaches evidenced in the eighth-grade Eight-Dimension-Cue group was on a par with that of the eighth-grade Four-Dimension-Cue and Eight-Dimension-No-Cue groups. However, regardless of whether irrelevant hypotheses were sampled, eighth-grade children exposed to a stimulus complex containing irrelevant information and provided with a cue list showed less focusing and more dimension checking than any other eighth-grade group. The predominant non-strategic pattern evidenced in this group involved the dimension checking error category which accounted for a larger proportion of problems in this eighth-grade group than any other. Further, the proportion of peob-lems showing a dimension checking plan successfully concluded remained below that of the two eighth-grade Four-Dimension groups.

In contrast to its effect on eighth-grade children, providing a cue list with stimuli containing only relevant information (Four-dimension) improved third-grade performance in most respects. Although none of the differences between the two third-grade groups on simple hypothesis testing measures reached significance, they were all in favor of increased efficiency with the provision of a cue list. Moreover, the difference between the two age groups in retaining an hypothesis following positive feedback and testing hypotheses previously sampled and disconfirmed was not significant in this condition. In addition, the third-grade children provided with a cue with a four dimensional configuration solved significantly more problems than their peers exposed to the identical stimuli without the cue list.

Third-grade children in the Four-Dimension-Cue group were also more likely to show a strategic approach to problem solution than third-grade children in the Four-Dimension-No-Cue group. Dimension checking was the modal strategy of third-grade children exposed to four-dimensional stimuli with and without a cue. However, the provision of a cue resulted in an increase in the proportion of problems in this category, while it had little effect on the manifestation of the other two strategies. The cue list also resulted in fewer stereotypic sequences and unsystematic sequences.

When irrelevant information was included in the stimulus complex, third-grade children like eighth-grade children sampled mostly relevant hypotheses. However, there were more problems showing irrelevant hypotheses than problems on which no irrelevant hypothesis was manifested. On all other measures of simple hypothesis testing efficiency, third-graders' performances in the Eight-Dimension-No-Cue condition was very similar to their peer's performances in the Four-Dimension-No-Cue condition. However, children in this age group exposed to a stimulus complex containing irrelevant information solved fewer problems than their peers exposed to only relevant information.

On those problems in which no irrelevant hypotheses were sampled on the first three trials, there was no difference between the third-grade Four-Dimension-No-Cue and Eight-Dimension-No-Cue groups either in the proportion of focusing or dimension checking. When all problems are considered, the third-grade Eight-Dimension-No-Cue group showed a larger proportion of strategic approaches than their peers in the Four-Dimension-No-Cue group. The increase was primarily in the dimension checking category which accounted for the majority of sequences that included irrelevant hypotheses. Of the non-strategic approaches, unsystematic sequences, with and without error were reduced while dimension checking error increased. Despite the increase in dimension checking approaches, there was a decrease in the proportion of problems on which this plan was successfully concluded in the Eight-Dimension-No-Cue group.

The provision of a cue with the stimulus containing irrelevant information also resulted in third-grade children sampling fewer irrelevant hypotheses than they did when no cue list was provided. However, there were no differences on any of the measures of hypothesis testing efficiency as a function of providing a cue with this configuration. There was also no significant difference on the probability of solving a problem due to the presence of a cue list. Moreover the third-grade children provided with a cue with a relevant-irrelevant stimulus complex solved significantly fewer problems than their peers provided with a cue with stimuli that contained no irrelevant information.

When the distribution included only those problems in which no irrelevant hypotheses are manifested during the first three trials, the proportion of dimension checking sequences was greater with cue list provision than without and equal to that observed in the Four-Dimension-Cue group. The proportion of focusing was, however, lower than that in either of the two comparable conditions. Overall, third-grade children exposed to irrelevant information with a cue list showed fewer strategic approaches than their peers exposed to only relevant information with a cue. The primary reduction was in the proportion of focusing sequences. Thus, among the third-grade children the provision of a cue with the relevant-irrelevant configuration did not have the same salutary effect on problem solving efficiency it had when only relevant information was included in the configuration. Children in this age group showed less efficient methods of reaching solution with the cue than without, with no significant increase in the probability of reaching solution. In contrast, providing the cue with only relevant information resulted in both an increase in strategic approaches and a greater probability of solution.

Discussion

It seems reasonable to begin with an evaluation of the effects of using a verbal probe in this paradigm. Traditionally, blank-trial probes have been used to infer the individual hypotheses manifested in the course of solving four-dimensional concept identification problems (Eimas, 1969, 1970; Gholson et al., 1972, 1973; Gholson & McConville, 1974; Ingalls & Dickerson, 1969). Two studies have compared the relative effects of these two methods. Karpf and Levine (1971) with college student subjects and Phillips and Levine (1975) with second- and sixth-grade children.

It is difficult to make precise comparisons between the results obtained in this study using a verbal probe and the previous studies that used either blank-trial or verbal probes. In the first place, third-grade children were not included as subjects in any of the others and eight-grade children were included only in the Ingalls and Dickerson (blank-trial probe) study. Second, not all measures analyzed in this study are comparable to those of previous studies. A few comparisons with the Four-Dimension-No-Cue groups are possible, however, and are presented below.

In relation to hypothesis testing efficiency, the probability of maintaining an hypothesis following positive feedback was measured by Gholson et al. (1972, 1973), Gholson and Danziger (1975), Ingalls and Dickerson (1969) and Karpf and Levine (1971). The probabilities of .93 for third-grade children and .98 for eighth-grade children obtained in this study are in the range obtained in the aforementioned studies regardless of probe technique. An examination of the data suggests that although there may be small age differences within given studies, values range above .90 among subjects between second-grade and adult. This conclusion is further supported by examination of the data obtained from a more conservative measure of hypothesis retention following positive feedback; i.e., including only trials prior to the criterion run. These data were reported by Phillips and Levine (1975). They reported no probe effect with either age group. The probability obtained in the present study (.85) for third-grade children was slightly lower than either of Phillips second-grade groups (.87-.89), whereas the probability for eighth-grade children in this study (.95) was slightly higher than either of Phillips' sixth-grade groups (.92).

There are also data from a number of previous studies which permit comparisons of the probabilities of retaining disconfirmed hypotheses. Under standard conditions, using blank-trial probes, the range of probabilities for second- and fourth-grade children was .08 to .13 Gholson et al., 1972, Gholson & Danziger, 1975, Phillips, 1974). With verbal probes, Phillips second-grade children showed a probability of .06 while the probability for third-grade children in the present study was .05. The probabilities for sixth-grade, eighth-grade and adult subjects with blank-trial probes range from .01 to .07 (Gholson et al., 1972, 1973; Ingalls & Dickerson, 1969; Karpf & Levine, 1971). The comparable probabilities for verbal probe groups were .02 for both Phillips and Levine's sixth-grade group and Karpf and Levine's adults and .01 for eighth-grade children of the present study. It appears, then, that verbal probes slightly improve performance in the early elementary school years. This may be related to the tendency among some younger children to hold one hypothesis regardless of feedback (stimulus preference). Phillips and Levine (1975) found that this tendency was reduced in second-grade children when a verbal probe was used. The small difference among older children may be a consequence of variations in overall procedure.

The probability of manifesting locally consistent hypotheses has also been examined in previous research (Gholson et al., 1972, 1973; Gholson & Danziger, 1975; Phillips & Levine, 1975). These probabilities ranged from .74 to .89 among second- and fourth-grade children under the two probe techniques. The probability was .82 for the third-grade children in the present study, which is about midway within that range. The probabilities for sixth-grade children ranged from .87 to .94. The probability for eighth-grade children in this study (.93) is within that range but slightly below that of adults with blank trial probes (.96). Thus, these probabilities do not appear to vary as a function of probe technique.

Another measure of hypothesis testing efficiency for which there are comparable data is the percentage of feedback trial responses that are consistent with the preceding hypothesis. With blank-trial probes, this type of response consistency for second- to sixth-grade children ranges from 90% to 93% (Gholson & Danziger, 1975; Phillips & Levine, 1975). The range for eighth-grade children and adults was 95% to 97% (Ingalls & Dickerson, 1969; Karpf & Levine, 1971). With a verbal probe, comparable percentages were 98 for second-grade children and 99 for sixth-grade children and adults (Karpf & Levine, 1971; Phillips & Levine, 1975). In the present study, both age groups were 100% response consistent. It is to be noted that there is no time lapse between manifestation (verbalization) of an hypothesis and response choice when the verbal probe technique is used. In both the Phillips study and the present one, children were trained to state verbally their hypotheses in the presence of the stimulus configuration but prior to the choice response. Therefore, choosing a configuration inconsistent with the verbally stated hypothesis could apparently result only from either an 'oops' error (Levine, 1966) or a misunderstanding of the relations between the manifested hypothesis and the visual stimulus array. The high percentage of response consistency shown with both probe techniques(at least 90%) suggests that by second-grade, children understand the relationship between the manifested hypothesis and the succeeding feedback trial choice. The nearly perfect response consistency found with the verbal probe at all ages suggests there may be less probability of an 'oops' error with this probe. With blank-trial probes, four different stimulus configurations (i.e., four trials) are used to infer an hypothesis. A response is consistent if the choice on the fifth trial (a feedback trial) is consistent with the hypothesis inferred from the previous four (blank-trial) choice responses. Thus, the complex procedure in the blank-trial probe task may lead to more 'oops' errors.

It appears, then, that the hypothesis testing behaviors of second-grade children through adults show little effect of the particular probe technique. With either blank-trial or verbal probes, the manifested behaviors appear to reflect similar underlying processes.

The proportion of response sequences in each hypothesis sampling system category has been reported by Gholson et al. (1972, 1973) and Gholson and Danziger (1975) for second-, fourth-, sixth- grade and adult subjects using blank-trial probes. Phillips and Levine (1975) have compared the performance of second- and sixth-grade children under blank-trial and verbal probe conditions. The Gholson studies have consistently found no age differences in four dimensional problems in the frequencies of occurrence of each system within the elementary school groups. Dimention checking was the modal strategy of second-, fourth-, and sixth-grade children, accounting for 50% to 60% of the problems in the distributions. The proportion of problems categorized as focusing did not exceed 15% in any of these age groups. Both of Phillips' second-grade groups showed a comparable proportion of problems in the focusing category in the first session of her study (eight problems). Her Blank Trial Probe group showed a somewhat smaller proportion of problems in the dimension checking category (36%) whereas the Verbal Probe group showed a slightly higher proportion (68%). However, both sixth-grade groups in the Phillips & Levine study showed a considerably larger proportion of focusing sequences than the Gholson et al.sixth-grade groups. In the blank-trial probe group, 32% of the problems were categorized as involving focusing and 56% involved dimension checking. In the verbal probe group, 40% of sixth-grade children's problems involved focusing and 45%, dimension checking. These latter findings were comparable to adult data obtained from blank-trial probes (focusing 45% to 50%, dimension checking 45% to 50%). This would suggest that sixth-grade children's performances under verbal probe conditions are closer to adult performances than sixth-grade children's performance under blank-trial probe conditions. However, Phillips' sixth-grade blank-trial probe group also showed a larger proportion of focusing sequences than was evidenced in the Gholson sixth-grade groups. Therefore, it is difficult to determine the extent to which the verbal probe facilitated the functioning of sixth-grade children in the Phillips' study. The proportion of focusing among third-grade children in the present study (20%) was higher than that observed among the Gholson et al.elementary school groups and lower than that observed in Phillips' sixth-graders under blank-trial probe conditions. The proportion of dimension checking was comparable to

Phillips' elementary school groups regardless of probe (45%). The distribution of eighth-grade problems in the present study was almost identical to that of Phillips' sixth-grade verbal probe group (focusing - 41%, dimension checking - 45%).

It is difficult to compare the proportional representation in the the other categories, hypothesis checking, stimulus preference and unsystematic in the present study with the proportions obtained by Gholson and Phillips, since the proportions in these categories are generally small (below 25%) in all studies and age groups. The proportions observed in the present study among children in both age groups were similarly small and within the ranges observed previously under both blank-trial and verbal probe procedures.

Comparisons of distributions of hypothesis sampling systems in the Gholson et al.and Phillips and Levine studies with the present one suggest that the proportion of problems of elementary school children categorized as focusing may be affected by specific experimental conditions. Generally, a larger proportion of problems of children in this age range are categorized as involving focusing under verbal probe than under blank-trial probe conditions. In the Phillips and Levine study, the proportion of focusing increased with practice (second session, additional 12 problems) among second-grade children in the verbal probe group, but not the blank trial probe group. The proportion of focusing among third-grade children in the present study exceeded that of all elementary school groups under blank-trial probe conditions except Phillips and Levine's sixth-grade group. Under verbal probe conditions, both sixth (Phillips and Levine) and eighth-grade (present study) children manifested bimodal distributions comparable to those observed among adults under blank-trial probe conditions.

Analysis of the task under the two probe conditions indicates that the verbal probe may simplify the task. It clearly facilitates the manifestation of a focusing strategy. It is quite possible that the extended interval between feedback trials in studies involving blank-trial probes interferes with the retention of information necessary to the manifestation of focusing. In addition, in the blank-trial task the orthogonal structure of the sequence of feedback stimuli may be obscured by the intervening sequence of blank trials which present a different organizational structure.

If, as in the present study, the purpose is to permit the subject to display his most sophisticated information processing behavior, the verbal probe appears to be the more efficient method of observing this behavior during a concept identification task. It presents the basic task without the clutter of extraneous stimulus configurations. In addition, it reduces the trial-by-trial memory demand to that which is essential for problem solution. One further advantage to the method is that proportionally more problems may be used in the systems analysis. Finally, procedurally, the verbal probe permits the collection of a greater amount of data per session.

There is, however, another consideration: Under verbal probe conditions, the subject is required to state a simple hypothesis prior to each choice response. When he is asked to state his best bet as to the solution, the assumption is that he should have one; i.e., he should be generating hypotheses. But since it may force the subject to generate hypotheses on each trial, it may not be the most <u>appropriate</u> technique for obtaining information concerning <u>whether</u> problem solution is always mediated by hypothesis testing (see Falmagne, 1970; Gholson & O'Connor, 1975). It may not, for example, be an appropriate method for examining the underlying problem solving processes of younger children or children suspected of cognitive disabilities. It is possible that younger children and children who have learning disabilities have difficulty integrating verbal symbols with visual arrays (Kendler & Kendler, 1962). Furthermore, it is possible that children and adults, under conditions in which a verbal hypothesis is not required, hold no hypothesis for several trials during problem solving; i.e., wait for information in a 'no hypothesis' state (Falmagne, 1970). This is an issue which will be examined in future research investigating the relative efficiency of blank-trial and verbal probes with kindergarten children and children with reading disabilities.

In this study, using verbal probes, there were significant differences between third- and eighth-grade children on each measure of problem solving behavior in the standard four dimensional problems (i.e., measures of simple hypothesis testing efficiency, sequential response patterns, and solution per se).

Of particular interest are those measures which had been predicted to differentiate the two age groups. Neimark (1971) has hypothesized that the development of techniques for systematic memory storage and retrieval might be a basic process underlying cognitive development. She found that third-grade children, in comparison to older children and adults, showed less organization in encoding material for recall and less awareness of the items that had been previously recalled. Based on that finding, it had been thought that the younger children in the present experimental situation would be more likely than the older to resample previously manifested and disconfirmed hypotheses. The confirmation of this hypothesis leads to the suggestion that developmental differences in memorization strategies may have played an important role in the performance of children of the two age groups of the present study. Some clarification of this issue is gained by comparing the performance of children of the two age groups in a situation in which a memory aid was provided with four dimensional problems and when it was not (Four-Dimension-Cue and No-Cue). Providing a cue list of the total set of potential solutions resulted in third-grade children showing fewer repetitions of hypotheses that had been previous disconfirmed. There was no difference among eighth-grade groups. They showed few instances of repeating disconfirmed hypotheses regardless of the experimental condition.

There is, however, an alternate explanation. Consistent with the findings of Ingalls and Dickerson (1969) and Gumer and Levine (1971), the error most associated with not solving among the children of both ages involved the omission of the solution dimension from the sampling set. It is reasonable to suggest that dimensions may be omitted more frequently than the observations from unsolved problems might suggest. Many problems are solved prior to the manifestation of hypotheses from all dimensions, so the data based on only unsolved problems may be a conservative estimate of this kind of error.

In the present study, this type of error was more frequent among the younger children, but provision of a cue list with four dimensional stimuli sharply reduced the number of problems in which it occurred. This suggests that the third-grader's propensity to sample disconfirmed hypotheses may be, at least, partially, a consequence of processing feedback in relation to a reduced set of dimensions. Further support for this contention is found in the difference between third-grade response sequences with and without the cue list. First, more problems were suitable for systems analysis when the cue was provided (80% vs 74%). Second, the largest decrease, with the provision of the cue, was in the unsystematic category (21% vs 11%) while the accompanying increase was in the dimension checking category (45% vs 56%). This change in sequential responding as a function of the cue suggests that many of the unsystematic sequences may result from working with a sampling set containing less than four dimensions. The finding that 42% of the problems in the No-Cue condition showed non-strategic approaches, while only 28% of the problems in the Cue condition did, lends additional support to this suggestion. Moreover, the major reduction was in the unsystematic error category. The proportion of problems in the dimension checking error category remained constant across conditions. It appears that without a cue, third-grade children had difficulty retaining the set of dimensions from which the solution hypothesis is drawn. When a memory aid was provided, it was used to monitor the set of dimensions. This resulted in fewer unsystematic response sequences, particularly error patterns, at the outset of the problem and less sampling of disconfirmed hypotheses throughout the problem.

These changes, as a function of providing a memory aid, were not observed in the problems of eighth-grade children. On the whole, providing a cue list to eighth-grade children had little effect on sequential responding. A slight difference in the direction of increased unsystematic approaches was observed when the cue list was provided (non-strategic approaches; Cue-22%, No-Cue-16%). The provision of this list appears to have interfered with systematic responding and led to errors in dimension checking. Perhaps, it acted as a 'distractor' in the sense that

Hagen and Hale (1973) have suggested 'imposed labeling' does for older children. Or it may have led to the perception of the task as a simple one that did not require a systematic procedure to reach solution. Recall that all groups were exposed to both four and eight dimensional problems in pretraining with no cue provided. Moreover, these children did solve as many problems with the cue as without, despite the increase in unsystematic sequences.

The question is what accounts for the younger children's difficulty in keeping track of the entire set. All the dimensions relevant to solution appear in the visual arrays on each trial, so there is no need to retain the set in memory. The only activity necessary is a search of the two configurations. There is ample evidence (e.g., local consistency data) in the present study as well as others (Gholson et al., 1972, 1973; Gholson & Danziger, 1975; Phillips, 1974) that children in this age group understand the relationship of feedback information to stimulus configurations and succeeding hypotheses. The search task must, therefore, be limited to the configuration designated by feedback as including the solution hypothesis. It has generally been assumed that third-grade children are capable of both differentiating integrated configurations into the dimensions involved and processing information in relation to the individual elements (Tighe & Tighe, 1969). It is not clear, therefore, why keeping track of the set of potential hypotheses was more difficult for the younger than the older children in this situation. Perhaps, remaining attentive to the individual elements over a sequence of trials in which values are presented in different combinations is difficult for younger children. The continuous need to separate the configuration and search for individual dimensional values may be a greater strain on the attentional processes of younger children than older. On the other hand, the difficulty may lie in memory processes. The omission of a dimension might also be a consequence of the imperfect memory organization hypothesized by Niemark (1971).

There is, however, another explanation for the differences observed in the two age groups: both in the frequency of omitted dimensions and the propensity to retest previously disconfirmed hypotheses. These may be related to differences in strategic approaches. As Bruner et al. (1956) have pointed out, different problem solving strategies make differential memory demands. Recall that eighth-grade children in the present study were bimodal in their strategic approach to this task, with approximately equal proportions of focusing and dimension checking manifested. But dimension checking was the modal strategy of third-grade children under all conditions, with little focusing.

Each of these strategies make differential attentional and mnemonic demands on the problem solver. In focusing, the dimensional values are differentiated at the outset of the problem or, at the least, following the first feedback trial. At that point, the four values on the positive configuration are encoded and retained in memory. Following he the second feedback, that set is retrieved. The positive configuration is searched for the values that match the retrieved set. Those two are reentered in memory. The process continues through the third feedback trial at which time the set of hypotheses is reduced to the solution. The task is complicated by the sequencing of feedback information. Following negative feedback, the working hypothesis is eliminated and attention must be redirected from the stimulus configuration of choice to the positive one. The memory task involves an initial encoding of four values and a systematic elimination at each trial, of half of the information in storage.

The attentional and mnemonic demands involved in a dimension checking plan are not as evident. There are two alternate 'strategies' that could result in the inference of a dimension checking plan. The assumption that a subject following a dimension checking plan initially encodes the potential solutions into pairs on each dimension may be taken to imply that he initially memorizes the total set of four dimensions. He may begin the problem by sampling one dimension and selecting one value of that dimension to test. If the consistent choice receives negative feedback, he eliminates that hypothesis and by logical inference, the dimensional opposite value. At that point, he retrieves the retained set of dimensions, samples one of them and searches the positive configuration for the locally consistent value on that dimension. This procedure continues until he reaches solution. This description is similar to a focusing plan, then, in that all the information necessary for solution is initially encoded and there is a systematic reduction in memory load. Dimension checking, of course, involves a slower process than focusing, since only one dimension is eliminated per negative feedback trial rather than half the memory load each trial. The attentional task is also similar to focusing, although simpler, since the search task occurs only on negative feedback trials and the match involves only one value each time.

The finding that 65% of all second and third hypotheses in dimension checking sequences were inconsistent with first trial feedback suggests that complete stimulus information is not encoded in memory in the beginning stages of dimension checking. It appears that the subject following a dimension checking plan may not fit the description outlined above in that he may not initially memorize the four dimensions. Instead, he may sample one dimension at a time but when he receives disconfirmation in the form of negative feedback, he searches the positive configuration for a new hypothesis representing an unsampled dimension. Thus, he identifies a new dimension following each negative feedback trial. This process involves more complex attentional and memory processes than the previous example. The search task involves an identification rather than a match. In order to avoid repeating dimensions, the previously sampled and disconfirmed dimensions must be retained and retrieved during each search. Moreover, the set of memorized dimensions is increased with each disconfirmation.

It, therefore, seems possible that the omission of dimensions and resampling of disconfirmed hypotheses result from aborted dimension checking plans of the latter type. The attentional process and memory organization necessitated by the use of that strategy seems to provide more opportunity for these kinds of errors than does focusing. It is possible, then, that the difference in the incidence of these errors among third- and eighth-grade children may be related to the differential use of a focusing approach, rather than a difference in either attentional skills per se or strategies specifically geared toward memorization.

With the provision of a memory aid, the incidence of sampling errors decreased and the manifestation of systematic response sequences increased among third-grade children without an appreciable increase in focusing. This finding lends support to the proposition that the strategy used reflects the underlying cognitive organization of the problem solver. Providing a memory aid to third-grade children did not increase the use of a strategy hypothesized to reflect a logical plan beyond the cognitive level of most third-grade children (see Gholson & McConville, 1974). Rather, a cue list appears to simply aid children in this age range in displaying their modal approach to this kind of task.

It might be argued that the higher incidence of dimension checking in the Cue condition obscured an increase in occurrence of actual focusing plans. Recall that the proportion of problems showing globally consistent patterns categorized as focusing is systematically reduced by a specific amount dependent on the number of dimension checking patterns evidenced in that feedback sequence (Levine, 1975, Appendix). Since the number of dimension checking sequences observed fluctuated with age and experimental condition, it is possible to question whether the statistical manipulations resulted in accurate psychological portraits of the problem solving strategies of children in this study.

In attempting to answer this question, each child was classified according to the dominant system he manifested. Since there were eight problems, an attempt was made to classify each child according to the system that accounted for the majority of his problems (five). No child in either of the Four-Dimension conditions showed five problems that were solely dimension checking or hypothesis checking so mixed patterns were identified that constituted an hierarchy of efficiency. The following seven categories were constructed to account for the protocols of all children.

1. Five globally consistent patterns solved following third feedback.
2. Five patterns showing either globally consistent patterns or dimension checking patterns which were solved with only one hypothesis per dimension manifested.
3. Five patterns showing either globally consistent or dimension checking patterns without regard to whether solution was reached without error.

4. Five strategic patterns which include hypothesis checking and dimension checking and/or globally consistent patterns.
5. Five patterns including dimension error patterns as well as dimension checking and/or globally consistent patterns.
6. Five error patterns, either all unsystematic or unsystematic and dimension checking error.
7. Miscellaneous - no five patterns fitting one of the above categories.

The number of children in each of the categories under each experimental condition is presented in Table 10. In the Four-Dimension groups, only three third-grade children show five globally consistent patterns either with or without the memory aid. The improvement in performance as a function of the memory aid is reflected in the increase in the number of children in the next two categories and the decrease in children in the last three categories. While only 12 children were in the first three categories when no memory aid was provided, 19 were in these categories with the provision of a cue list.

In contrast, the number of eighth-grade children in the Four-Dimension conditions in the first three categories was not affected by the cue list. There were 10 children in the first category and 21 in the first three categories in both conditions. Thus, the difference, as a function of age, in the number of children in the _first_ category, regardless of the availability of a cue list supports the contention, inferred from the frequency distribution of the systems analysis, that the use of a focusing plan (i.e.,responding in a globally consistent fashion) is beyond the cognitive level of most third-grade children.

Since it was originally hypothesized that the use of a focusing plan may be dependent on the logical operations available, this finding

Table 10

Number of Children Classified in Each Category for Each Experimental Group

	Categories						
Third grade	1	2	3	4	5	6	7
Four-Dimension-No-Cue	2	3	6	2	3	5	2
Four-Dimension-Cue	3	6	10	2	1	1	1
Eight-Dimension-No-Cue	0	0	12	4	6	1	1
Eight-Dimension-Cue	0	2	7	4	6	3	2
Eighth grade							
Four-Dimension-No-Cue	10	8	3	3	0	0	0
Four-Dimension-Cue	10	8	3	1	1	1	0
Eight-Dimension-No-Cue	4	3	11	2	4	0	0
Eight-Dimension-Cue	6	6	5	3	3	0	1

was not unexpected. Nevertheless, it is possible that the attentional and mnemonic demands imposed by the different strategic approaches dictated the choice of a particular strategy rather than that the initial strategic plan dictated the kinds of attentional and mnemonic skills that would be used. Unfortunately, this cannot be determined from the data of the present study.

Further research addressed to this issue will examine the problem solving approaches of children in these age ranges in a situation in which concrete replicae of the set of hypotheses are available. The subjects can then choose which hypotheses they will retain and discard during the course of a problem. This procedure will permit the observation of how the child approaches the problem as well as the attentional and mnemonic skills he uses in his attempt to find the solution.

The original impetus for the present study was Hagen's findings of age related changes in processing relevant and irrelevant information. In this study, the relevant and irrelevant information were presented in the context of a complex concept identification task rather than the simple recall task used in the Hagen studies. Consistent with Hagen's findings, children in both age groups used irrelevant information and third-grade children were more likely to do so than eighth-grade children. Nevertheless, considerably less than half of the hypotheses generated by children in either age group were from irrelevant dimensions. This suggests that contrary to the original hypothesis the younger, as well as the older, children made an effort to encode the relevant set. The children in the older group were, however, more successful.

The more interesting question in this study concerned how the presence of irrelevant information affected the problem solving approaches among children of the two age groups. It appears that the most crucial aspect of the task in the Eight-Dimension-No-Cue condition was the initial encoding of the set of relevant hypotheses. Recall that the relevant dimensions were only identified at the outset of the problem. In addition the total set of dimensions remained the same throughout the eight problems whereas the designation of the relevant subset changed from problem to problem. Therefore, the subject had to commit a new set of four dimensions to memory for each problem and retain that set throughout the problem. During that period, a sequence of visual arrays was presented in which the relevant set of dimensions was differentially integrated with four irrelevant dimensions. Thus, remembering which four dimensions were relevant for each problem may have constituted a complex task in itself. When those memory demands were met, there was very little difference in the kinds of strategies manifested by children of either age group from that used by their peers who were exposed to visual arrays containing only relevant information. If no irrelevant hypotheses were manifested in the first three trials, for example, the proportion of focusing among the children of each age was similar to that evidenced in the Four-Dimension-No-Cue condition.

It was originally thought that eighth-grade children would show a smaller proportion of focusing and a larger proportion of dimension checking when irrelevant information was included in the stimulus configuration. The reasoning was that children in this age group would recognize the difficulties inherent in the task when irrelevant information was included. Consequently, they were expected to attempt a strategy that entailed less 'cognitive strain.' However, an analysis of the demands of the task in this situation suggests that focusing might have involved less 'cognitive strain' than dimension checking. The exclusion of irrelevant information from sampling depends on initiating the problem by encoding the four relevant dimensions. This is assumed to be an integral and necessary part of a focusing plan. Although the attentional task is more complex when irrelevant values must be excluded from processing during the course of the problem, there is no difference in the memory task as a function of the presence of irrelevant information. However, following a dimension checking plan is much more difficult when irrelevant information is present than when it is not. In order to exclude irrelevant dimensions from sampling, the four relevant dimensions must be encoded in memory. Following each negative feedback trial, the unsampled dimensions must be retrieved and the positive configuration, containing both relevant and irrelevant values, must be searched for a value from a previously unsampled but relevant dimension. Thus both the memory and attention aspects of the task are complicated by the presence of irrelevant information. In this situation, then, the necessity of excluding irrelevant hypotheses from sampling dictates that the four relevant dimensions should be encoded and retained at the outset of the problem, regardless of the strategy followed. In addition, the attentional task may be almost as complicated in dimension checking as in focusing but the memory load is reduced at a slower rate in dimension checking than focusing. This suggests that the focusing plan may involve less 'cognitive strain' than a dimension checking plan in this situation. Thus, if a focusing approach were within the cognitive capacity of a child, we might expect him to use it.

Therefore, with hindsight, it is not surprising that eighth-grade children manifested a higher proportion of focusing than dimension checking sequences on those problems in which they retained the set of relevant dimensions. It is more surprising that, overall, there was a reduction in the use of a focusing approach by eighth-grade children when the problems contained irrelevant dimensions. Only four eighth-grade children in contrast to ten in each of the Four-Dimension groups showed a focusing approach on the majority of problems (see Table 10). It is not clear what underlies this reduction. Possibly, in the Four-Dimension conditions, the memory aspect of the problem begins with encoding the set of hypotheses on the first positive configuration. The necessity of encoding the four relevant dimensions prior to first feedback in the Eight-Dimension-condition then increases the memory aspect of the task. In addition, there is more likelihood of error in processing present stimulus information in relation to the retrieved information when irrelevant dimensions are included in the stimulus. Although dimension checking was the modal strategy of children in this age group when irrelevant hypotheses were manifested, there was no difference in the proportion of dimension checking sequences observed in this group and the eighth-grade Four-Dimension-No-Cue group. Instead, the reduction in focusing sequences was accompanied by an increase in error sequences. Thus it does not seem as if eighth-graders altered their strategic approach when irrelevant information was included in the stimulus array. Rather, the reduction in focusing shown by this eighth-grade group in relation to the eighth-grade Four-Dimension groups appears to reflect an increase in the propensity to commit mnemonic and attentional errors in processing information in this situation.

Regardless of whether irrelevant dimensions were included in the sampling set, the modal strategy of third-graders was dimension checking.

When irrelevant information was included in the stimulus array, a larger proportion of the problems of third-grade children showed a strategic approach, primarily dimension checking, than when no irrelevant information was present. It appears that the inclusion of irrelevant dimensions provided third-grade children with a more conducive situation in which to demonstrate their preferred strategy. The inclusion of extra dimensions enabled children in this age group to display their modal strategy on a larger number of problems if the irrelevance of half of the sampling set was ignored. However, there were fewer problems solved by this third-grade group than any other.

Although eighth-grade children showed more regression from their optimal level of performance in this condition than third-grade children did, they maintained their superiority over the younger group in all aspects of problem solving: they were more likely to exclude irrelevant hypotheses from sampling, and they were more likely to show a strategic approach, particularly focusing. It is difficult to determine what underlies this superiority. It may be that the older children are more proficient in encoding and retaining relevant information. It may be that they are more proficient in excluding irrelevant information from intruding in the problem solving process. On the other hand, the superior performance may reflect a more advanced cognitive organization which showed some deterioration in this situation but still enabled eighth-grade children to cope with this problem in a more efficient fashion than third-grade children.

In comparing the present results with Hagen's work, it must be noted that there are important differences between the two experimental situations. This study not only employed a more complex task but it also involved a stimulus situation in which the separation of relevant and irrelevant information was probably more difficult. Addressing themselves to this issue, Hagen and Hale (1973) have delineated the processes in the two stages of Neisser's (1967) theory of information processing in relation to developmental differences in selective attention. According to them, the first stage involves identification of relevant information. The second stage involves maintenance of attention to relevant information while ignoring irrelevant information. They hypothesized that developmental differences are found in the second stage following discrimination of relevant from irrelevant information. In situations in which first stage processing involves considerable effort (i.e., the initial discrimination between relevant and irrelevant information is difficult), there is less likelihood of developmental differences in performance.

In the present study, the second stage represented a particularly arduous task as maintenance of attention to the relevant information included retaining identification of that set in memory. This was complicated by combining relevant and irrelevant information in an integrated configuration (See Hagen & Hale, 1973 for a discussion of this issue). This made first stage processing (i.e., discrimination of the relevant from the irrelevant) a continuous necessity throughout the sequence of trials in each problem until solution was reached. Additionally, the difficulty of first stage processing was compounded by the alternation of particular dimensions as relevant or irrelevant across problems. Thus, the small difference between performances of third- and eighth-grade children in this situation in which relevant and irrelevant were combined in each stimulus configuration is consistent with the Hagen and Hale thesis. It is left for future research to determine if larger differences in problem solving efficiency would be observed in these two age groups if the relevant and irrelevant elements were spatially separated and the identification of each as relevant or irrelevant were constant as they typically were in the Hagen studies.

When the configuration contained irrelevant information, the cue list was provided specifically to examine the selection and exclusion process when the necessity of memorizing the relevant set of dimensions was eliminated. Children in both age groups sampled fewer irrelevant hypotheses when the memory aid was provided. This suggests that the large proportion of irrelevant hypotheses generated when the memory aid was excluded was related to difficulty in encoding and retaining the relevant hypotheses set during the problem. Even with the provision of a cue list, however, children of both age groups sampled irrelevant hypotheses. This suggests that excluding irrelevant information from processing entailed more than just memorization,since it was presumably not required when the cue list was provided. The fact that age differences remained, even with the provision of the cue list, indicates that there are developmental differences in the exclusion process as well as in the encoding and retention processes.

For third-grade children, providing a memory aid with stimuli that included irrelevant information appears to have had mixed effects. Fewer irrelevant hypotheses were manifested when the cue list was provided; but fewer strategic approaches were evidenced with the cue; and there was no increase in the probability of solution related to presence of the cue list. Furthermore, there was a smaller proportion of focusing in this condition than in any other with no accompanying increase in dimension checking. More error sequences were in evidence when the eight-dimensional stimuli were presented with a cue list than when that configuration was presented without one. One may speculate that the effort to use the cue list in selecting hypotheses may have interfered with the generation of more efficient sequential plans.

The memory aid (cue list) did not have the same beneficial effect on third-grade performance when the stimulus configuration contained irrelevant information as it had when it accompanied a configuration containing no irrelevant information. It appears that using the cue list to generate hypotheses when the list of hypotheses matched the elements in the visual array presented a very different situation than when the list had to be used to select elements embedded in a configuration which also contained elements to be excluded. Two explanations which are not mutually exclusive are suggested. It is possible that the deployment of attention between the cue list and the visual array in a situation in which the search task involves both separation and exclusion interfered with the efficient use of the cue list. It is also possible that the memory aid does not affect the primary difficulty in a stimulus situation of this kind i.e., excluding the visual irrelevant elements from intrusion on the processing of relevant elements.

The eighth-grade children made more effective use of the cue list in this situation than the third-grade children. In contrast to their peers exposed to the relevant-irrelevant configuration without a cue list, the eighth-grade children provided with a cue list did trade off focusing approaches for dimension checking approaches. This leads to the speculation that children in this age group conceived of the exclusion aspect of the task as most crucial when the configuration included irrelevant information but the need to encode and retain the relevant set was eliminated. If that was the case, the reduction in focusing and concomitant increase in dimension checking may be because the older children have an hierarchy of strategies available. The cue list may have been perceived as a concrete prop which reduced the memory aspect of the task. Dimension checking, then, may have been considered by some eighth-grade children an approach more likely to lead to success in this situation than focusing, since it involves less demanding selective attention. The likelihood of error in the visual search aspect of the task would seem to be greatly reduced if only one hypothesis per trial has to be identified from a list of hypotheses. Therefore, the increased use of dimension checking in this situation may be because older children attempt to analyze both the task and their goals in an attempt to employ the most efficient strategy in the particular situation. Whether or not it was a consequence of the alternation in strategic approach, the provision of a memory aid did result in a larger proportion of successes in the eight dimensional situation than in the identical situation when no memory aid was provided.

Nevertheless, the presence of irrelevant information in the visual array did interfere with the efficiency of children in this age group. While they solved as many problems as their peers exposed to configurations that included only relevant information, they were more likely to show dimension checking errors and less likely to reach solution using that approach.

Eighth-grade children showed more efficiency than third-grade children in both the approach to and completion of problems in this situation. Notwithstanding, the finding that there was less efficiency among eighth-grade children when irrelevant information was included in the stimulus complex with the cue list, than in either condition without irrelevant information suggests that the exclusion aspect of the task required effort even for children of this age. It is not possible to determine if this difficulty was a consequence of difficulty in first stage processing as postulated by Hagen and Hale to be inherent in the presentation of relevant and irrelevant information as an integrated configuration. It is also not possible to determine whether the developmental improvement in offsetting irrelevant information inferred from the Hagen (1967, 1970) studies to occur at older ages than those of this study would be evident among older subjects in stimulus situations in which first stage processing was difficult since related research using older subjects has not yet been reported.

Nevertheless, the present study indicates that selective attention requires more than just appropriate encoding strategies. Despite the fact that children of both age groups were able to use the cue list as an aid in generating almost exclusively relevant hypotheses, the irrelevant information in the stimulus configuration affected problem solving efficiency in both age groups.

In regard to this, it is to be noted that the primary question of the present study differed from that of the Hagen studies in an important way. The Hagen studies address themselves to the question of whether irrelevant information is attended to and remembered in a simple memory task. The primary concern here was how the presence of irrelevant information affects the processing of relevant information in a problem solving task and how children of different ages handle this situation.

There was considerable variation among the eighth-grade children's performances in response to the different experimental conditions. It appears that children in this age range analyze a situation and adjust their method to the particular stimulus situation. This is possible because they have a repertoire of information processing techniques that can be applied in different situations. In contrast, the younger children may have a smaller repertoire of such techniques available. It is, of course, not clear if their failure to adjust their processing in response to different conditions is solely a function of the limited range of techniques available to them or stems from failure to accurately assess the stimulus situation.

In this study, the eighth-grade children were, like adults in previous research (Gholson et al.,1972, 1973), equally likely to use a focusing or dimension checking strategy to solve problems when the stimuli contained only relevant information. When the task was complicated by the inclusion of irrelevant information and the necessity of retaining the relevant information in memory, they appeared to concentrate on the retention of the relevant information. The difference in the shape of the category frequency distribution in the Eight-Dimension-No-Cue condition when problems manifesting irrelevant hypotheses are excluded and included suggests that success in retaining the relevant information was likely to lead to the use of a focusing strategy which minimized the memory demands of the task. If the relevant information was not retained, a dimension checking strategy was more likely to be followed. When the retention of the relevant set was not necessary due to provision of a cue list, they appeared to concentrate on the exclusion aspect of the task. The finding that they generated fewer irrelevant hypotheses with the cue list than without suggests they may have used the memory aid in following a dimension checking plan which simplified the selective attention aspect of the task.

Few third-grade children appear to have the cognitive resources needed to manifest a focusing strategy. Therefore, they were limited to a dimension checking approach regardless of the stimulus condition. The manifestation of this approach appeared to depend on whether the stimulus situation provided an opportunity to display it. The younger children may have been less aware than the older of the set of dimensions that were potential solutions to the problem. This would help explain why more successful completions of problems were evidenced when the stimuli contained only relevant information and a cue list was provided than in any other condition. The provision of an external aid that represented exactly the elements in the stimulus configuration appears to have greatly increased the probability of successfully concluding a dimension plan in this age group.

In this study the manifestation of different strategies has been identified as possibly representing different conceptual levels of development. One question of major concern is whether these strategies are unique to these particular tasks. The identification of consistent sequential response patterns within individual children permits the investigation of developmental differences in logical planning and processing of information to be carried out across a number of problem solving tasks. Strategies involved in other tasks may indicate similar differences in conceptualization associated with developmental levels. For example, although the twenty questions task differs from a concept identification task in form as well as content, the strategies identified in that task, constraint seeking and hypothesis seeking appear to parellel focusing and dimension checking or hypothesis checking. The strategies identified by Bruner et al. (1956) also appear to parallel focusing and dimension checking. Neimark and Lewis, (1967, 1968) too, have identified strategic approaches to a diagnostic problem solving task which follow a similar developmental pattern to that observed in this and related studies. It might be expected that a child who manifests a focusing system in concept identification tasks would show a similar strategy in a concept attainment task and would manifest constraint seeking in the twenty question task. In response to Neimark and Lewis' (1967,1968) diagnostic problem solving task, that same child might show the general logical principles that characterize the approach of adolescents. If consistent developmental differences are involved in the manifestation of different problem solving strategies, they should be evidenced across various types of problems. Future research will be addressed to this question.

Appendix A

Task demands in the use of strategies

Hypothesis checking - Each attribute of each dimension is methodically tested and eliminated, dimension by dimension. New hypotheses are consistent with current information but not necessarily with any information received on past feedback trials. There are no logical inferences necessary. The use of such a strategy implies that the problem solver is able to examine a stimulus configuration and recognize that it exists on a continuum of different dimensions and can be compared with and discriminated from another configuration on that basis - e.g. it is larger, its shape is triangular, not circular etc. Moreover, he must understand the instructions that the problem is soluble only if the configuration is broken up; that one of the hypotheses on one of the dimensions is the solution. He then must attend to both sets of stimuli on each trial and relate them to feedback information. This strategy makes the least demands on the subjects' intellectual resources and the largest on his mnemonic resources (encoding, storing and retrieving). To reach solution without repeating hypotheses he must continuously recall the hypothesis he has tested and the memory load increases as the problem progresses.

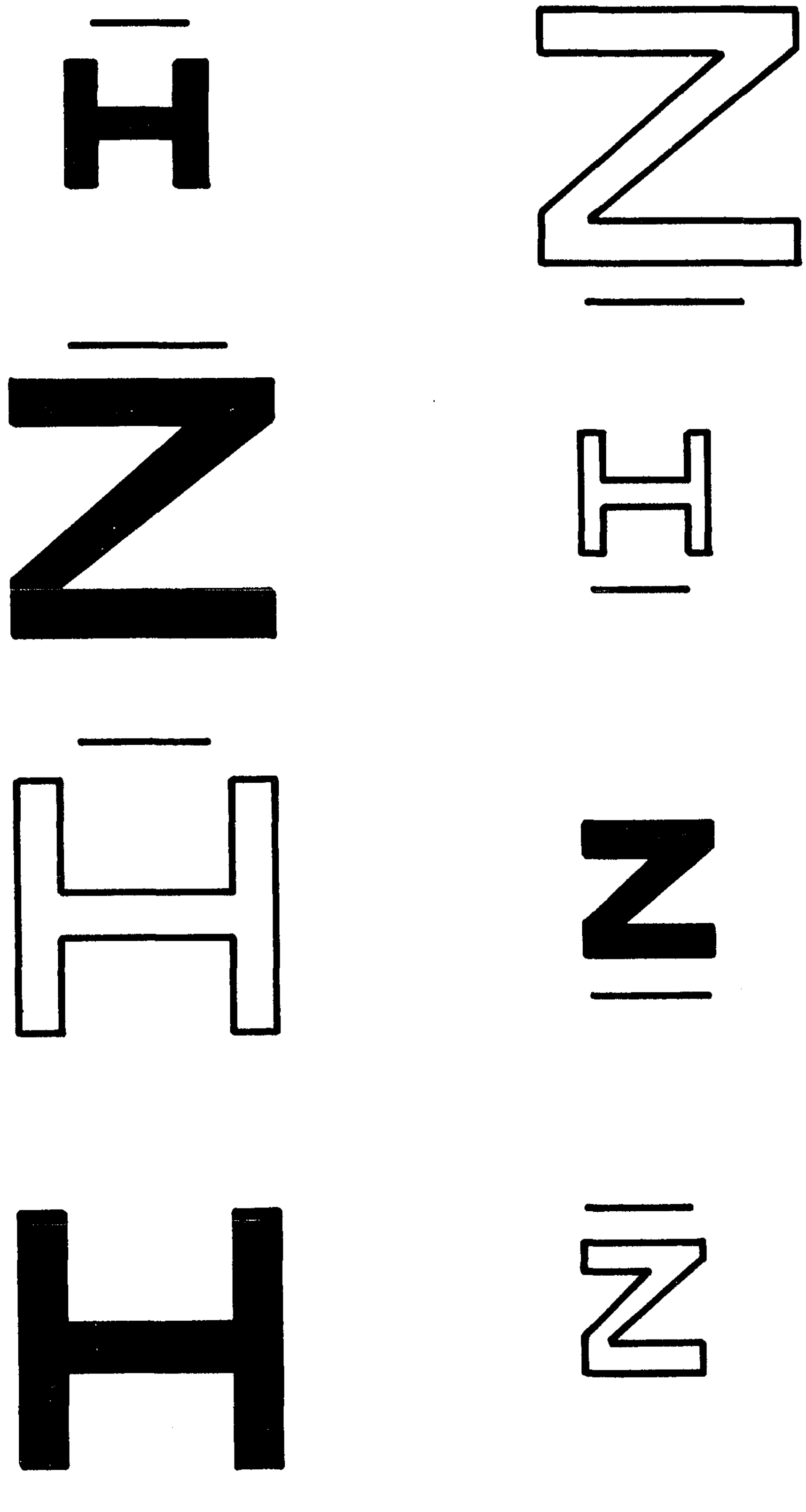
Dimension checking - This is a more sophisticated version of Hypothesis checking. The use of such a strategy seems to imply operational thinking. To use it efficiently (errorlessly) entails a systematic ordering of the dimensions into four mutually exclusive sets, each comprised of a pair of complementary values. Information on one of the values is used to process the set. The subject must infer that if an hypothesis appeared on the positive configuration on the previous feedback trial, the dimensionally opposite hypothesis must have appeared on the negative.

Therefore, if the first hypothesis is now on the negative picture, neither can be the solution. Thus this strategy necessitates the use of logical reasoning in the elimination of hypotheses. In a sense, each dimensional set is viewed as a superordinate class with two subordinate members. If information is received that one member is potentially correct, the possibility of the other being correct is automatically precluded by the complementary nature of the membership in the class. The reasoning would seem to be of this order; color = black, white; if black, not white, if not black, not color. Thus the class is eliminated as a potential solution on the basis of the relationship between the members of the class to each other and to the total set. The use of such logic implies the ability to join objects in a class and to understand their relations; to simultaneously view either as a separate subordinate member of the class and both as the total superordinate class. Using this strategy reduces the memory load during successive trials. The set of hypotheses has logically been reduced to four. Thus, the most hypotheses a subject must keep mnemonic track of is three.

Focusing seems to involve complex operations. The task is begun with a long range plan in which all hypotheses are considered as the potential solution. At each step of the problem, the set is reduced by half. Following the first feedback information, there are four potential solutions which must be encoded in some symbolic fashion, stored and retrieved following the second feedback. The retrieved set must be compared to the currently positively designated set and a new set formed from the common elements of both sets. Repeating this operation following third feedback leaves a remainder set of one- the solution. Solution is achieved by means of a successive series of subtractive or cancelling out operations. The initiating of such a plan would appear to involve logical analysis of the problem structure as well as the ability to engage in sequential elimination operations. The Flavell (1963) description of the formal operational thinker suggests that that stage of thinking is reflected in the use of this strategy.

Appendix B

Example of orthogonality



Appendix C

Instructions

Preliminary I - 4 Dimension Stimuli

I would like you to help me with some puzzles. What will happen is that you will see pictures like these. As you can see, there are two pictures. One has an 'X' and one has a 'T', one has a 'circle' and one has a 'square', and one the letter is colored in so its 'full' and the other is 'empty', one letter is 'big' and one is 'little'. (point while naming).

One of these is the right answer - and the right answer is right for the whole puzzle. So, for example if 'big' is the right answer, this picture (point) would be right since the letter is'big'and this other picture would be wrong since the letter is 'little' (point) What you'll be doing is choosing one of the two pictures, the one you think is right. If you think this picture is right (point) then point here---if you think this picture is right, point here (point). After you've picked a picture, I'll tell you whether the one you chose was right. Try to be right as often as you can. If you figure out the right answer---What is always right, then you can pick the right picture every time.

Now let's try this puzzle. Remember (pointing) what is right all the time could only be 'big' or 'little', 'full' or 'empty', all the 'Xs' or all the 'Ts' or the 'circles' or the 'squares'. Your job is to figure out what the right answer is - remember its only one of those and its right all the time. (Feedback is given after each choice.)

AFTER TRIAL 6 Did you notice what was right all the time.

IF CORRECT Yes, 'full' was the right answer this time. That was very good.

IF WRONG IF COMPOUND ____ No only one is always right, did you notice which one was always right.

IF WRONG_____No, that wasn't right all the time.

IF WRONG OR DOESN'T KNOW BY 6TH: EITHER/OR HINT. I'm going to give you a hint about the right answer. What will always be right here is either the '<u>full</u>' one or the '<u>empty</u>' one (point).

AT END : Did you figure out what was right every single time?

IF YES: Very good,'<u>full</u>'was the right answer.

IF WRONG: The right answer was '<u>full</u>' let's try it again. (Show first pair and point to 'full') I want you to pick every picture with a letter thats 'full' and see what happens. (Say 'right' after each choice.)

CRITERION: TEN IN A ROW: See, now that you know the right answer, you're right all of the time.

AFTER: This is the way all of the puzzles will go. The right answer will always be <u>one</u> of the things on one of the two pictures you see each time. Like this time it was the 'full' letter but it could have been the 'X' or the 'T', the empty letter or the'circle' or the 'square' or the 'big' one or the 'little' one.

Do you have any questions?

<u>Preliminary 2</u> - 8 Dimension Stimuli

OK let's try this next puzzle. In this puzzle, there are lots of things but I'll tell you the only things the right answer might be and it will <u>only be one</u> of the things I tell you - so listen and look at what I point at. It could be the 'single border' or the 'double border', it could be 'green' or it could be 'orange', or it could be the 'dashed border'or the 'solid border' or it could be that the 'line is on top' of the picture or 'on the bottom' of the picture. Your job again is to figure out which <u>one</u> is the right answer - <u>whats right</u> <u>every</u> <u>single</u> <u>time</u>.

AFTER TRIAL 6: What was always right?

IF CORRECT: Yes, that was very good 'solid border' was the right answer.

IF INCORRECT AND RELEVANT: GIVE EITHER/OR HINT AND THEN TELL

IF INCORRECT AND IRRELEVANT OR COMPOUND: No, remember, I told you it could only be 'a single border' or a 'double border, a 'dashed border' or a 'solid border', 'orange' or 'green', or the'line on the top' or the 'line on the bottom'.

GIVE EITHER/OR HINT AND THEN TELL

AT END: See, once you get the right answer, you are right all the time.

Preliminary 3 - 4 Dimensional Stimuli

Let's try this next puzzle. In this one the right answer could be 'square' or 'circle', 'big' or 'little', 'top' or 'bottom', 'dashed' or 'solid'. (pointing to them) One of these is right all of the time and it s your job to figure out what it is.

AFTER TRIAL 3: Now tell me what your best bet is about the right answer- the one that's right all of the time. (From Trial 3 on, ask child - what's your best bet about the right answer?)

AFTER TRIAL 6: What was right all of the time (Big)

IF CORRECT OR INCORRECT PROCEDURE SAME AS IN 1 and 2.

Preliminary 4 - 8 Dimensional Stimuli

In this puzzle, the right answer could be 'big' or 'little', double line' or 'single line','purple' or 'red' or 'top' or 'bottom'. Remember these are the only possible right answers.

BEFORE PROBLEM WITH 1st SET IN FRONT OF CHILD - What do you think the right answer will be this time?

AFTER VERBALIZATION: Show me the picture that's right.

AFTER FEEDBACK: What's your <u>best bet now</u> about what's the right answer?

AFTER 6th FEEDBACK: What I want you to do from now on is before each time you pick a picture, tell me what's <u>your best bet</u> about the right answer - which you think it might be. If you forget to tell me, I'll remind you. I want you to tell me just as you've been doing when I've asked but I just don't want to have to ask you each time.

REMIND IF CHILD DOESN'T

GIVE EITHER/OR HINT OR TELL LIKE OTHER PRELIMINARIES.

Preliminary 5

Ok let's look at these, this time the answer is either 'big' or 'little', 'empty' or 'full', 'blue' or 'brown', 'V' or 'A'. Remember these are <u>all</u> the possible answers. <u>One</u> of them is the right answer. It's your job to figure out which it is. Remember before you choose a picture, tell me your best bet about what the right answer is. If you don't have a best bet at the beginning - tell me that - but after the first choice always tell me which one you think the answer is.

General Instructions

BEFORE PROBLEM: Don't forget to tell me <u>before</u> you choose a picture what your best bet is about what's the right answer.

AFTER LISTING THE ATTRIBUTES: The answer is <u>only</u> one of those.

BEFORE PROBLEM: AFTER TELLING RELEVANT ATTRIBUTES: Now you tell them to me. Tell me what could be the answer. Point to them.

IF WRONG OR FORGETS: NO, the answer is either REPEAT.

AT END OF PROBLEM: That was very good. That was a good try.

Appendix D

Sample cue card

X	R
Circle	Square
Big	Little
Dashed Border	Solid Border

Appendix E

The Analysis of Variance Summary Table

for the Probability of Solution

Source	SS	DF	MS	F
Age	7122.986	1	7122.986	65.23*
Information	6264.441	1	6264.441	57.365*
Cue	2048.918	1	2048.918	18.762*
Age x Information	49.217	1	49.217	-
Age x Cue	45.60	1	45.60	-
Cue x Information	293.115	1	293.115	2.68
Age x Cue x Information	309.398	1	309.398	2.83
Within	20093.367	184	109.203	

*P .05 = 3.91

Tukey b

3rd 8 DNC	8th 8 DNC	3rd 8 DC	3rd 4 DNC	3rd 4 DC	8th 8 DC	8th 4 DNC	8th 4 DC
46.38	53.23	53.82	56.71	64.30	66.55	13.09	13.98

$P < .05$

The Analysis of Variance Summary Table for the Probability that the Same Hypothesis was Maintained on Successive Trials When the Intervening Feedback was Positive

Source	SS	DF	MS	F
Age	1340.9	1	1340.9	29.26*
Information	413.5	1	413.5	9.02*
Cue	2.0	1	2.0	.04
Age x Information	65.12	1	65.12	1.42
Age x Cue	128.95	1	128.95	2.81
Cue x Information	45.0	1	45.0	.98
Age x Cue x Information	3.48	1	3.48	.08
Within	843.3	184	45.83	

Tukey *b*

3rd 8 DNC	3rd 4 DNC	3rd 8 DC	8th 8 DC	3rd 4 DC	8th 8 DNC	8th 4 DC	8th 4 DNC
75.66	77.11	77.18	79.13	79.28	81.96	84.62	84.67

Analysis of Variance Summary Table for the Probability that the Same Hypothesis was Maintained on Successive Trials When the Intervening Feedback was Positive Excluding H_o and Criterion Run Data

Source	SS	DF	MS	F
Age	887.6068	1	887.6068	11.8048*
Information	586.8114	1	586.8114	7.8044*
Cue	3.7416	1	3.7416	
Age x Information	68.4249	1	68.4249	
Age x Cue	296.5094	1	296.5094	3.9438*
Cue x Information	202.6222	1	202.6222	2.6948
Age x Cue x Information	40.5912	1	40.5912	
Within	13834.9214	184	75.1897	

Tukey *b*

3rd 8 DNC	3rd 4 DNC	8th 8 DC	3rd 8 DC	3rd 4 DC	8th 8 DNC	8th 4 DNC	8th 4 DC
69.10	70.27	70.43	70.73	74.17	75.61	77.33	78.10

The Analysis of Variance Summary Table for the Probability that a Disconfirmed Hypothesis was Maintained

Source	SS	DF	MS	F
Age	338.8578	1	338.8578	12.1906*
Information	2.1357	1	2.1357	
Cue	20.2085	1	20.2085	
Age x Information	13.9699	1	13.9699	
Age x Cue	.4360	1	.4360	
Cue x Information	9.2477	1	9.2477	
Age x Cue x Information	6.9767	1	6.9767	
Within	5114.5599	184	27.7965	

The Analysis of Variance Summary Table for the Probability of Generating a Locally Consistent Hypothesis Following Negative Feedback

Source	SS	DF	MS	F
Age	1428.17	1	1428.1736	30.19*
Information	1021.8072	1	1021.8072	21.60*
Cue	.0732	1	.0735	
Age x Information	91.3421	1	91.3421	1.93
Age x Cue	.01	1	.01	
Cue x Information	31.2565	1	31.2565	
Age x Information x Cue	82.9362	1	82.9362	1.75
Within	8705.7786	184	47.3140	

Tukey *b*

3rd 8 DC	3rd 8 DNC	3rd 4 DNC	8th 8 DNC	3rd 4 DC	8th 8 DC	8th 4 DC	8th 4 DNC
62.92	65.09	66.2	67.84	68.27	68.32	73.80	74.30

The Analysis of Variance Summary Table for the Probability

of Retesting a Previously disconfirmed Hypothesis

Source	SS	DF	MS	F
Age	3227.3560	1	3227.3560	44.26*
Information	211.7200	1	211.7200	2.90
Cue	324.2720	1	324.2720	4.45
Age x Information	27.4993	1	27.4993	
Age x Cue	65.7072	1	65.7072	
Cue x Information	14.1071	1	14.1071	
Age x Information x Cue	159.7948	1	159.7948	2.19
Within	13418.1079	184	72.944	

Tukey *b*

8th 4 DC	8th 4 DNC	8th 8 DC	8th 8 DNC	3rd 4 DC	3rd 8 DC	3rd 8 DNC	3rd 8 DC
13.69	13.84	15.26	17.98	19.65	23.36	24.76	25.76

The Analysis of Variance Summary Table for the Probability of Generating an Irrelevant Hypothesis

Source	SS	DF	MS	F
Age	599.35017	1	599.35017	8.16*
Cue	4219.4646	1	4219.4646	57.48*
Age x Cue	153.4445	1	153.4445	2.09
Within	6754.0540	92	6754.0540	

The Analysis of Variance Summary Tables for the Percentage of Problems Showing Irrelevant Hypotheses

Source	SS	DF	MS	F
Age	2829.8988	1	2829.8988	14.26*
Cue	9258.6888	1	9258.6888	46.66*
Age x Cue	257.6115	1	257.6115	1.30
Within	18255.0935	92	198.42	

The Analysis of Variance Summary Table for the Number of Focusing Plans (Based on Three Randomized Observations)

Source	SS	DF	MS	F
Age	608.03	1	608.03	34.57*
Information	403.44	1	403.44	22.94*
Cue	26.88	1	26.88	1.53
Age x Information	20.54	1	20.54	1.11
Age x Cue	19.44	1	19.44	1.11
Information x Cue	44.83	1	44.83	2.55
Age x Information x Cue	.35	1	.35	-
Within	281.41	16	17.59	-

Tukey b

3rd 8 DC	3rd 8 DNC	3rd 4 DNC	8th 8 DC	3rd 4 DC	8th 8 DNC	8th 4 DC	8th 4 DNC
2.9	6.2	9.6	9.6	12.3	16.0	22.1	23.6

The Analysis of Variance Summary Table for the Number of Dimension Checking (Based on Three Randomized Observations)

Source	SS	DF	MS	F
Age	8.4	1	8.4	
Information	181.49	1	181.49	6.63*
Cue	206.50	1	206.50	12.91*
Age x Information	26.4	1	26.4	
Age x Cue	228.3	1	22.8	
Cue x Information	60.18	1	60.18	3.76
Age x Cye x Information	13.74	1	23.14	
Within	437.91	16	27.37	

Tukey *b*

3rd 8 DNC	8th 8 DNC	3rd 8 DC	3rd 4 DNC	8th 4 DC	8th 4 DNC	8th 8 DC	3rd 4 DC
12.0	15.5	21.1	21.1	22.7	23.8	24.4	27.5

The Analysis of Variance Summary Table for the Percentage

of Problems Solved Strategically

<u>Source</u>	<u>SS</u>	<u>DF</u>	<u>MS</u>	<u>F</u>
Age	8316.3308	1	8316.3308	60.4042*
Information	4703.8860	1	4703.8860	34.1658
Cue	95.6828	1	95.6828	
Age x Information	45.4546	1	45.4546	
Age x Cue	72.6928	1	72.6928	
Cue x Information	65.3566	1	65.3566	
Age x Cue x Information	522. 2582	1	522.2582	3.7933
Within	25332.7646	184	137.6780	

Tukey <u>b</u>

3rd 8 DC	3rd 8 DNC	3rd 4 DNC	8th 8 DNC	3rd 4 DC	8th 8 DC	8th 4 DC	8th 4 DNC
32	33.82	38.29	43.95	45.39	46.26	55.0	56.95

Appendix F

Chi Square

Probability Problem Order Affected Solution

	Third Grade	Eighth Grade		
Four-Dimension-No-Cue	$X^2_{(3)}$ 4.16	$X^2_{(3)}$ 2.95		
Four-Dimension-Cue	$X^2_{(3)}$ 7.84*	$X^2_{(3)}$ 1.53		
Eight-Dimension-No-Cue	$X^2_{(3)}$ 2.98	$X^2_{(3)}$ 2.32		
Eight-Dimension-Cue	$X^2_{(3)}$ 1.76	$X^2_{(3)}$ 6.18	.10	.20

$X^2_{(3)}$ = 7.815 p .05

Overall X^2 (24) = 29.72 p .20

$X^2_{(24)}$ = 36.42

Number of Problems Solved as a Function of Problem Order

Condition	Total Solved	1	2	3	4	5	6	7	8	Number Errors	Problem Number
Third Grade											
Four-Dimension-No-Cue	129	12	19	18	16	15	12	18	19	2	1,1
Four-Dimension-Cue	153	17	19	15	18	22	18	23	21	3	1,3,7
Eight-Dimension-No-Cue	101	12	11	11	9	14	15	16	13	2	1,1
Eight-Dimension-Cue	125	18	17	11	16	20	13	13	17	2	3,7
Eighth Grade											
Four-Dimension-No-Cue	176	22	21	23	20	23	24	22	21	1	4
Four-Dimension-Cue	179	22	20	24	22	22	23	23	23	1	1
Eight-Dimension-No-Cue	132	18	14	17	16	15	17	17	18	0	-
Eight-Dimension-Cue	162	16	20	21	19	20	23	22	21	0	-

Appendix G

Hypothesis Sampling Systems Calculations
(From Gholson et al, 1972)

Four-Dimension-No-Cue

Third Grade

Feedback Pattern	Focusing	Dimension Checking	Hypothesis Checking	Stimulus Preference	Unsystematic	Tot
+ + -	14-8 6	16+8 24	1	3	0	34
+ + -	18-3.6 14.4	6+3.6 9.6	2	0	14	40
- - -	11-6.6 4.4	11+6.6 17.6	6	1	8	37
- + -	8-4.5 3.5	8+4 12	1+1+.5 2.5	3	8	29
Total	28.3	63.2	11.5	7	30	140
Percent	20.2	45.1	8.2	5.0	21.4	73.7

Eighth Grade

Feedback Pattern	Focusing	Dimension Checking	Hypothesis Checking	Stimulus Preference	Unsystematic	Tot
+ + -	28-7 21	14+7 21	0	0	0	42
+ - -	21-6.6 14.4	11+6.6 17.6	1	0	8	41
- - -	21-8.4 12.6	14+8.4 22.4	7	0	2	44
- + -	25-5.5 19.5	9+4.5 13.5	0+2+1 3	0	1	37
Total	67.5	74.5	11	0	11	164
Percent	41.2	45.4	6.7	0	6.7	85.9

Hypothesis Sampling Systems Calculations
(From Gholson et al, 1972)

Four-Dimension-Cue

Feedback Pattern	Focusing	Dimension Checking	Hypothesis Checking	Stimulus Preference	Unsystematic	Total
			Third Grade			
+ + -	20-8.5 11.5	17+8.5 25.5	2	1	0	40
+ - -	17-10.2 6.8	17+10.2 27.2	1	0	8	43
- - -	14-8.4 5.6	14+8.4 22.4	6	0	7	41
- + -	12-6 6	7+3.5 10.5	2+5+2.5 9.5	1	2	29
Total	29.9	85.6	18.5	2	17	153
Percent	19.5	55.9	12.1	1.3	11.1	80.1
			Eighth Grade			
+ + -	29-6 23	12+6 18	0	0	0	41
+ - -	26-4.8 21.2	8+4.8 12.8	1	0	8	43
- - -	15-8.4 6.6	14+8.4 22.4	5	0	7	41
- + -	20-5.5 14.5	10+5 15	1+1+.5 2.5	0	0	32
Total	65.3	68.2	8.5	0	15	157
Percent	41.6	43.4	5.4	0	9.6	81.8

Hypothesis Sampling Systems Calculations
(From Gholson et al, 1972)

Eight-Dimension-No-Cue
(Irrelevant hypotheses excluded)

Third Grade

Feedback Pattern	Focusing	Dimension Checking	Hypothesis Checking	Stimulus Preference	Unsystematic	Total
+ + -	12-2 10	4+2 6	3	5	0	24
+ - -	5-5.4 0	9+5.4 14.4	1	0	4	19(.4)
- - -	6-3 3	5+3 8	5	0	2	18
- + -	7-2.5 4.5	4+2 6	0+1+.5 1.5	3	0	15
Total	17.5	34.4	10.5	8	6	76 (.4)
Percent	22.9	45.1	13.7	10.5	7.9	39.8

Eighth Grade

Feedback Pattern	Focusing	Dimension Checking	Hypothesis Checking	Stimulus Preference	Unsystematic	Total
+ + -	19-5 14	10+5 15	2	1	0	32
+ - -	20-3.6 16.4	6+3.6 9.6	0	0	4	30
- - -	9-4.2 4.8	7+4.2 11.2	9	0	5	30
- + -	18-3 15	4+2 6	1+2+1 4	0	0	25
Total	50.2	41.9	15	1	9	117
Percent	42.9	35.7	12.8	.8	7.7	60.9

Hypotheses Sampling Systems Calculations
(From Gholson et al, 1972)

Eight-Dimension-Cue
(Irrelevant hypotheses Excluded)

Third Grade

Feedback Pattern	Focusing	Dimension Checking	Hypothesis Checking	Stimulus Preference	Unsystematic	Total
+ + -	12-7.5 4.5	15+7.5 22.5	3	3	0	33
+ - -	4-4.8 0	8+4.8 12.8	1	0	13	26(.8)
— — —	1-6 0	10+6 16	6	0	6	23(5)
- + -	9-6.5 2.5	7+3.5 10.5	1+6+3 10	0	1	24
Total	7	61.8	20	3	20	106(+5.8)
Percent	6.3	55.3	17.9	27	17.9	57.8

Eighth Grade

Feedback Pattern	Focusing	Dimension Checking	Hypothesis Checking	Stimulus Preference	Unsystematic	Total
+ + -	13-4 9	8+4 12	3	0	0	24
+ - -	15-8.4 6.6	14+8.4 22.4	2	0	2	33
— — —	11-9 2	15+9 24	5	0	4	35
- + -	17-6.5 10.5	9+4.5 13.5	0+4+2 6	0	1	31
Total	28.1	71.9	16	0	7	123
Percent	23.1	58.9	13.1	0	5.7	64.1

Hypotheses Sampling Systems Calculations
(From Gholson et al, 1972)

Eight-Dimension-No-Cue
(Irrelevant hypotheses added)

Third Grade

Feedback Pattern	Focusing	Dimension Checking	Hypothesis Checking	Stimulus Preference	Unsystematic	Total
+ + - Irrelevant	12-2 10	4+2 +12 18	3	5	0	36
+ - - Irrelevant	5-5.4 0	9+5.4 +15 29.4	1	0	4	34(.4)
— — — Irrelevant	6-3 3	5+3 +7 15	5 + 5 10	0	2 +5 7	35
- + - Irrelevant	7-2.5 4.5	4+2 +13 19	1+.5 +1 2.5	3	0	29
Total	17.5	81.4	16.5	8	11	134(.4)
Percent	13.0	60.6	12.3	6.0	8.2	70.2

Eighth Grade

Feedback Pattern	Focusing	Dimension Checking	Hypothesis Checking	Stimulus Preference	Unsystematic	Total
+ + - Irrelevant	19-5 14	10+5 +8 23	2	1	0	40
+ - - Irrelevant	20-3.6 16.4	6+3.6 +10 19.6	0	0	4	40
— — — Irrelevant	9-4.2 4.8	7+4.2 +8 19.2	9 +1 10	0	5	39
- + - Irrelevant	18-3 15	4+2 +9 15	1+2+1 +1 5	0	0	35
Total	50.2	76.8	17	1	9	154
Percent	32.6	49.9	11.0	.6	5.8	80.7

Hypotheses Sampling Systems Calculations
(From Gholson et al, 1972)

Eight-Dimension-Cue
(Irrelevant hypotheses added)

Third Grade

Feedback Pattern	Focusing	Dimension Checking	Hypothesis Checking	Stimulus Preference	Unsystematic	Total
+ + - Irrelevant	12-7.5 4.5	15+7.5 +8 30.5	3	3	0	41
+ - - Irrelevant	4-4.8 0	8+4.8 +9 21.8	1	0	13	35(.8)
- - - Irrelevant	1-6 0	10+6 +6 22	6 +3 9	0	6 +2 8	34(5)
- + - Irrelevant	9-6.5 2.5	7+3.5 +2 12.5	1+6+3 10	0 +2 2	1 1	28
Total	7.0	86.8	23	5	22	138(5.8)
Percent	4.9	60.4	16.0	3.5	15.3	74.5

Eighth Grade

Feedback Pattern	Focusing	Dimension Checking	Hypothesis Checking	Stimulus Preference	Unsystematic	Total
+ + - Irrelevant	13-4 9	8+4 +5 17	3	0	0	29
+ - - Irrelevant	15-8.4 6.6	14+8.4 +3 25.4	2	0	2 +1 3	37
- - - Irrelevant	11-9 2	15+9 +3 27	5	0	4	38
- + - Irrelevant	17-6.5 10.5	9+4.5 +1 14.5	4+2 6	0	1	32
Total	28.1	83.9	16	0	8	136
Percent	20.7	61.7	11.8	0	5.9	70.8

Hypotheses Sampling Systems Calculations
(All problems)

Four-Dimension-No Cue

Third Grade

F.P.	Fo	Fo Sh	D-Ch	D-Ch Sh	H-Ch	SP	Unsyst	D-Ch Error	Unsyst Error	Total
+ + -	12-9 7	+4	18+9 28	+1	1	3	0	4	4	4
+ - -	18-3.6 14.4		6+3.6 9.6		2	0	8	5	10	49
- - -	11-6.6 4.4		11+6.6 17.6		6	1	2	7	10	48
- + -	8-4.5 3.5		8+4 15	+3	1+1+.5 2.5	3	7	9	6	46
Total	29.3		70.2		11.5	7	17	25	30	190
Percent	15.4		36.9		6.1	3.7	8.9	13.2	15.8	

Eighth Grade

F.P.	Fo	Fo Sh	D-Ch	D-Ch Sh	H-Ch	SP	Unsyst	D-Ch Error	Unsyst Error	Total
+ + -	28-7 24	+3	14+7 22	+1	0	0	0	1	0	47
+ - -	21-6.6 14.4		11+6.6 17.6		1	0	5	6	4	48
- - -	21-8.4 12.6		14+8.4 22.4		7	0	2	4	1	49
- + -	25-5.5+2		9+4.5 13.5		1+2+1 4	0	1	6	1	47
Total	72.5		75.5		12	0	8	17	6	191
Percent	38.0		39.5		6.3	0	4.2	8.9	3.1	

Note:
F.P= Feedback Pattern
Fo= Focusing
Fo Sh=Focusing with shift
D-Ch= Dimension Checking
D-Ch Sh= Dimension Checking with shift
H-Ch= Hypothesis Checking
SP=Stimulus Preference
Unsyst= Unsystematic
D-Ch Error=Dimension Checking with error
Unsyst Error=Unsystematic with error

Hypotheses Sampling Systems Calculations
(All problems)

Four-Dimension-Cue

Third Grade

F.P.	Fo	Fo Sh	D-Ch	D-Ch Sh	H-Ch	SP	Unsyst	D-Ch Error	Unsyst Error	Tota
+ + -	20-8.5+2 13.5		17+8.5 26.5	+1	2	1	0	3	2	48
+ - -	17-10.2 6.8		17+10.2 27.2		1	0	5	3	5	48
- - -	15-8.4 6.6		14+8.4 22.4		5	0	4	6	4	48
- + -	10-7 7	+4	9+4.5	+1	2+5+2.5 9.5	1	2	11	2	47
Total	33.9		90.6		17.5	2	11	23	13	191
Percent	17.7		47.4		9.2	1.0	5.8	12.0	6.8	

Eighth Grade

+ + -	29-6 25	+2	12+6 18		0	0	0	5	0	48
+ - -	26-4.8 21.2		8+4.8 12.8		1	0	8	2	3	48
- - -	15-8.4 6.6		14+8.4 22.4		5	0	6	7	1	48
- + -	20-5.5+1 15.5		10+5 19	+4 +4	1+1+.5 2.5	0	0	9	2	48
Total	68.3		72.2		8.5	0	14	23	6	192
Percent	35.6		37.6		4.4	0	7.3	12.0	3.1	

Hypotheses Sampling Systems Calculations
(All problems)
Eight-Dimension-No-Cue

Third Grade

F.P.	Fo	Fo Sh	D-Ch	D-Ch Sh	H-Ch	SP	Unsyst	D-Ch Error	Unsyst Error	Total
+ + -	11-2.5	+2	5+2.5	+2	3	5	0	6	2	
Irrel			+12							
	10.5		21.5							48
+ - -	5-5		9+5.4(-.4)		1	0	2	12	4	
Irrel			+15							
	0		29							48
- - -	5-3.6		6+3.6		5	0	5	9	6	
Irrel			+7		+5					
	1.4		16.6		10					48
- + -	7-2.5	+1	4+2	+3	1+1+.5	3	0	8	5	
Irrel			+13		+1					
	5.5		22		3.5					47
Total	17.4		39.1		17.5	8	7	35	17	191
Percent	9.1		46.6		9.2	4.2	3.7	18.3	8.9	

Eighth Grade

F.P.	Fo	Fo Sh	D-Ch	D-Ch Sh	H-Ch	SP	Unsyst	D-Ch Error	Unsyst Error	Total
+ + -	19-5	+1	10+5		2	1	0	3	2	46
Irrel			+8							
	15		23							
+ - -	20-3.6		6+3.6		0	0	3	8	2	49
Irrel			+10							
	16.4		19.6							
- - -	9-4.2		7+4.2		9	0	3	6	5	48
Irrel			+8		+1					
	4.8		19.2		10					
- + -	18-3		4+2	+1	1+2+1	0	0	10	3	49
Irrel			+9		+1					
			16		5					
Total	51.2		77.8		17	1	6	27	12	192
Percent	26.7		40.5		8.9	.5	3.1	14.1	6.3	

<u>Hypotheses Sampling Systems Calculations</u>
(All problems)

<u>Eight-Dimension-Cue</u>

<u>Third Grade</u>

F.P.	Fo	Fo Sh	D-Ch	D-Ch Sh	H-Ch	SP	Unsyst	D-Ch Error	Unsyst Error	Total
+ + - Irrel	11-8 5	+2	16+8 +8 32		3	3	0	4	1	48
+ - - Irrel	4-4 0		8+4.8 (-.8) +9 21		1	0	6	13	7	48
- - - Irrel	1-1 0		10+6 (-5) +6 17		6 +3 9	0	3	10	9	48
- + - Irrel	7-7.5 0		9+4.5(-.5)+2 +2 17		1+6+3 10	0 2 2	1	12	6	48
Total	5		87		23	5	10	39	23	192
Percent	2.6		45.3		12	2.6	5.2	20.3	12	

<u>Eighth Grade</u>

F.P.	Fo	Fo Sh	D-Ch	D-Ch Sh	H-Ch	SP	Unsyst	D-Ch Error	Unsyst Error	Total
+ + - Irrel	13-4 +8 17		8+4 +5 19	+2	3	0	0	9	0	48
+ - - Irrel	15-8.4 6.6		14+8.4 +3 25.4		2	0	2 +1 3	9	2	48
- - - Irrel	11-9 2		15+9 +3 27		5	0	1	9	4	48
- + - Irrel	17-6.5+4 14.5		9+4.5 +1 16.5	+2	1+4+2 7	0	1	8	1	48
Total	40.1		87.9		17	0	5	35	7	192
Percent	20.9		45.8		8.9		2.6	18.2	3.6	

Appendix H

(From Gholson, Levine and Phillips, 1972)

The technique for analyzing a set of protocols into the various systems will be demonstrated here.

The general conception: It is assumed that a subject begins a problem with some system that dictates his mode of sampling hypotheses. Each system is manifested in one of a small subset of hypothesis sequences. For example, the Stimulus Preference System is manifested when the same hypothesis is continually repeated after each of a series of disconfirmations. There is a small subset of these hypothesis sequences since the particular hypothesis is not specified -- a subject might perserverate on "large size," or on "black," etc.

A second assumption is that this system is maintained at least until the fourth feedback trial, i.e., the hypotheses observed after each of the first three feedback trials are dictated by a single system.

The third assumption concerns the subjects' manner of resampling for hypotheses after a disconfirmation. The hypotheses are divided into two classes: Response-set hypotheses (defined as the subject's preference, or bias, causing him to produce long strings of stereotypic responses) and Prediction hypotheses (defined as the subject's prediction concerning the solution). Correspondingly, the systems are divided into two classes: Stereotypes (e.g., Stimulus Preference), in which the subject persists in the same hypothesis despite repeated disconfirmations, and Strategies, in which the subject follows some plan which, in principle, leads to solution. During Strategies an hypothesis is genuinely tested, i.e., is rejected following a disconfirmation. The third assumption is that when the subject rejects an hypothesis and resamples (i.e., during Strategies) he always selects a locally-consistent hypothesis.

Finally, the set of systems is specified _a priori_. The following set was considered:

Stimulus Preference (abb. S-P): The subject selects an hypothesis and persists with it for the three feedback trials.

Hypothesis Checking (abb. _H_-Ch): The eight hypotheses are ordered by the subject into the pairs of hypotheses from each of the four dimensions, as though the subject imagines a list of the four pairs of hypotheses. He goes through this list, testing each hypothesis in turn, one dimension at a time. Thus, he tries an hypothesis, then, if it is disconfirmed, he tries its complement (the opposite hypothesis on the same dimension). If that is disconfirmed he tries an hypothesis from another dimension, then its complement, etc.

Dimension Checking (abb. D) : As with _H_-Ch the hypotheses are ordered by the four dimensions. In this case, the subject recognizes that after the first trial disconfirmation of an hypothesis also logically disconfirms the entire dimension. (This is logically correct since the just-disconfirmed hypothesis was, when it had been sampled, locally consistent. Its complement, therefore, of necessity had at that time been locally inconsistent). After each disconfirmation, therefore, the hypothesis comes from a different dimension.

Focusing (abb. Fo): With each feedback the subject eliminates all disconfirmed hypotheses, whether explicitly manifested or not. Thus, each hypothesis manifested is consistent with all the preceding feedback-trial information. The hypothesis after the third feedback trial is, of necessity, the solution.

The analysis: In principle, the experimenter's decision about the system employed by a particular subject on a particular problem is simple. He inspects the first three hypotheses and, on the basis of the particular hypothesis sequence, decides which system is being manifested. However, a more detailed treatment of the general analytical technique is required because there are certain small complications. There is an occasional overlap in the manifestation of two systems. For example, a subject who is following the D system may, when resampling after the third feedback trial, select the correct dimension and, perforce, the solution hypothesis. The resulting hypothesis sequence might be identical to that of a subject following the Fo system. Such confounding requires special techniques for arriving at valid estimates of the frequency with which the various systems occur. These techniques vary somewhat with each type of confounding and with the feedback sequence.

Because of these complications the categorizing of protocols will be described separately for each sequence of feedback trials. Problems with confounding will be discussed as they arise. Before proceeding, however, a few definitions will be helpful: (1) the feedback, "right" or "wrong" will be symbolized by + or -, respectively. A subscript (e.g., $+_2$) will indicate the trial on which the feedback occurred. (2) The subscripts i, j, k, and l will refer to the four dimensions. Thus, the symbol $\underline{H}_j$ will refer to the first hypothesis seen from the j^{th} dimension. If the j^{th} dimension is color and "black" is the first hypothesis observed from this dimension, the $\underline{H}_j$ = black. The complementary, or opposite, hypothesis will be indicated by the prefix "op." Thus, if $\underline{H}_j$ = black, then op$\underline{H}_j$ = white. In general, the subscripts i, ---, l will be used in chronological sequence, so that _i_ will be employed for the first dimension manifested in a problem, _j_ for the second, etc. (3) The correct H, i.e., the solution, will be symbolized as H^+. (4) Corresponding to the Sy labels, S-P, H-Ch, D, and FO, will be the frequencies of these Sys: f(S-P), f(H-Ch), etc.

Sequence: $+_1 \; -_2 \; -_3$

Each experimental problem is tentatively catalogued as one of the four Sys according to the manifestation indicated.

Stimulus Preference (S-P) manifestation: $+_1 H_i \quad -_2 H_i \quad -_3 H_i$

Hypothesis Checking (H-Ch) manifestation: $+_1 H_i \quad -_2 op H_i \quad -_3 H_j$

Dimension Checking (Tentative cataloguing, indicated by priming the symbol, D') manifestation: $+_1 H_i \quad -_2 H_j \quad -_3 (H_k = H^+)$

Focusing (tentative cataloguing, indicated by priming the symbol, Fo') manifestation: $+_1 H_i \quad -_2 H_j \quad -_3 (H_k = H^+)$. Also, H_j must be not only locally consistent (see assumption 3, above) but consistent with the first feedback trial as well. If this criterion is not met then the sequence is interpreted as a manifestation of D'.

Correction required: After cataloguing the problems according to the sets of manifestations noted above, a correction is required. An S following the D Sy may choose an H_j which is consistent with the first feedback trial and an $H_k = H^+$. The resulting H sequence would mistakenly be catalogued as Fo thereby inflating its frequency estimate, f(Fo), and correspondingly decreasing f(D). The relative frequency with which this particular H sequence occurs (given that S holds D and responds so as to receive the feedback sequence $+_1 \; -_2 \; -_3$) may be estimated from the following considerations: The corrected D value contains two components, those sequences which cannot be confused with the Fo sequence (labelled D') and those sequences which resemble the Fo sequence (labelled D*). The frequency of D is the sum of the two frequencies, i.e.,

$$f(D) - f(D') + f(D^*),$$

where f(D') is obtained directly from the data. Similarly, the sequences in the category labelled Fo' contain two different sets of sequences: those which are actually the result of the Fo system (labelled Fo) and those which are from the D system (D*). These two frequencies combine giving

$$f(Fo') \text{ and } f(Fo) - f(D^*),$$

where f(Fo') is obtained directly from the data.

These two equations contain three unknowns. A third equation is obtained from these considerations: With orthogonal stimulus sequences a subject who is holding the D system and who will receive $+_1$ and $-_2$ may start the problem with one of only two hypotheses (there are four Hs which lead to a correct response on trial 1; two of these lead to an error on trial 2). Starting with one of these two hypotheses, D*, the sequence resembling the Fo sequence, will occur one-fourth of the time; starting with the other hypothesis that Fo-resembling sequence will occur one-half the time. Since the probability that this subject (holding the D system) will start the problem with each of these two hypotheses is one-half, the probability of D* occurring is given by

1/2(1/4) + 1/2(1/2)= 3/8. Similarly, D' will occur 5/8 of the time. Therefore, the requires third equation is

$$f(D^*) = (3/5)f(D').$$

Solving the three equations yields:

$$f(D) = f(D') + (3/5)f(D')$$

$$\text{and} \quad f(Fo) = f(Fo') - (3/5f(D').$$

Thus, the value (3/5)f(D') must be added to and subtracted from, respectively, the initial, tentative frequency assignments for the D ststem and the Fo system

Sequence: $-_1$ $-_2$ $-_3$

S-P manifestation: $-\underline{H}_{1-i}$ $-\underline{H}_{2-i}$ $-\underline{H}_{3-i}$

H-Ch manifestation: This system has two sets of sequences depending upon whether one assumes that the subject holds an hypothesis before the problem starts or whether one assumes that the subject receives information from the first trial before applying his system. Each procedure yields its own unique sequence. Either sequence, therefore, was regarded as an instance of H-Ch.

These sequences are: $-\underline{H}_{1-i}$ $-\underline{H}_{2-j}$ $-_3 op\underline{H}_j$ or $-\underline{H}_{1-i}$ $-_2 op\underline{H}_{-i}$ $-\underline{H}_{3-j}$

D' manifestation: $-\underline{H}_{1-i}$ $-\underline{H}_{2-j}$ $-_3(H_{-k} \neq \underline{H}^+)$

The same confounding between D and Fo exists with this feedback sequence as with the preceding. Also, the same correction procedure is appropriate:

f(D) = f(D') + (3/5)f(D'), and

f(Fo) = f(fo') - (3/5)f(D'),

where f(D') and f(Fo') are frequencies obtained from the initial cataloguing:

Sequence: $+_1$ $+_2$ $-_3$

S-P manifestation: $+\underline{H}_{1-i}$ $+\underline{H}_{2-i}$ $-\underline{H}_{3-i}$

H-Ch manifestations: $+\underline{H}_{1-i}$ $+\underline{H}_{2-i}$ $-_3 op\underline{H}_i$

D' manifestation: $+\underline{H}_{1-i}$ $+\underline{H}_{2-i}$ $-_3(\underline{H}_j \neq \underline{H}^+)$

Fo' manifestation: $+\underline{H}_{1-i}$ $+\underline{H}_{2-i}$ $-_3(\underline{H}_j = \underline{H}^+)$

Here, again, a subject holding the D system would, if the correct dimension were second on his list, produce a sequence resembling focusing. Since the correct dimension cannot be first on this subject's list (the given feedback sequence has an error on trial 3, an impossible event if the subject had started out holding $\underline{H}^+$) it may be either second, third, or fourth. One-third of the time it will be second and a Fo sequence will result, i.e.,

$P(D^*/D) = 1/3$, and $P(D'/D) = 2/3$.

Therefore, $f(D^*) = 1/2)f(D')$ and

$$F(D) = f(D') + (1/2)f(D')$$

$$f(Fo) = f(Fo') - (1/2)f(D')$$

Sequence: $-_1\ +_2\ -_3$

S-P Manifestation: $\underset{1-i}{-\underline{H}}\ \underset{2-i}{+\underline{H}}\ \underset{3-i}{-\underline{H}}$

H-Ch' Manifestation: As with the sequence $-_1\ -_2\ -_3$, it is necessary to recognize that H-Ch can have one of two forms: the subject may hold an hypothesis before the first feedback trial or he may start through his list after receiving information from the first feedback trial. In the latter case no difficulties arise. The manifestation is $\underset{1-i}{-\underline{H}}\ \underset{2-i}{+\underline{H}}\ -_3 op\underline{H}_i$. If, however, the subject uses the initial feedback trial to reject the first hypothesis from the first dimension on his list, then the disconfirmation at trial 3 causes a shift to a new dimension, yielding the sequence $\underset{1-i}{-\underline{H}}\ \underset{2-i}{+\underline{H}}\ \underset{3-j}{-\underline{H}}$. It should be obvious that this sequence is identical to that produced by D and, if $\underline{H}_j = \underline{H}^+$, by Fo. The confounding with D will be treated here; the confounding with Fo will be treated subsequently.

I, $\underline{H}_j = H^+$ then the sequence could be the result of H-Ch or of D.

A simple way to deconfound is to view the next hypothesis the subject employs. If his system is H-Ch, then he will show $-H_{1-i} \; +\underline{H}_{2-i} \; -\underline{H}_{3-j} \; -_4 op\underline{H}_j$. This clearly is not a product of D. If, on the other hand, the subject is holding the D system, then he will show $-\underline{H}_{1-i} \; +\underline{H}_{2-i} \; -\underline{H}_{3-j} \; -\underline{H}_{4-k}$.

In the summary, (H-Ch') manifestation:

$(H\text{-}Ch)_1$: $-\underline{H}_{1-i} \; +\underline{H}_{2-i} \; -_3 op\underline{H}_i$, or

$(H\text{-}Ch')_2$: $-\underline{H}_{1-i} \; +\underline{H}_{2-i} \; -_3(\underline{H}_j \neq H^+) \; - \underline{H}_{4-k}$

Fo' menifestation: $-\underline{H}_{1-i} \; +\underline{H}_{2-i} \; -_3(\underline{H}_j = \underline{H}^+)$

Correction required: Whereas with previous feedback sequences only the D and Fo systems yielded the same hypothesis sequence, here three systems, $(H\text{-}Ch)_2$ as well as D and Fo, can yield the same hypothesis derivation in the previous cases.

Thus, the corrected frequency for $(H\text{-}Ch)_2$ contains two components, those sequences, here labelled $(H\text{-}Ch')_2$, not confuseable with the other two system patterns, and the sequence which resembles that from the Fo system, labelled $(H\text{-}Ch^*)_2$. The frequency of $(H\text{-}Ch)_2$ is the sum of the two frequencies, i.e.,

$$f(H\text{-}Ch)_2 = f(H\text{-}Ch')_2 + f(H\text{-}Ch^*)_2.$$

Similarly, $f(D) = f(D') + f(D^*)$

Also, $f(Fo) + f(Fo') - f(H\text{-}Ch^*)_2 - f(D^*)$.

Values for the primed variables are obtained directly from the data. In order to determine the corrected values of the systems, i.e., $f(H\text{-}Ch)_2$, $f(D)$ and $f(Fo)$, it is necessary first to determine the values of $f(H\text{-}Ch^*)_2$ and $f(D^*)$. These may be obtained from the following considerations:

For $(H\text{-}Ch)_2$: A subject's hypothesis sequence will resemble the Fo system only if the correct dimension is second on his list. Since it cannot be first on the list (H_1 cannot be H^+ because of the error given at the third feedback trial), it will be second one-third of the time. Therefore, $P\left[(H\text{-}Ch^*)_2 \mid (H\text{-}Ch)_2\right] = 1/3$, and $P\left[(H\text{-}Ch')_2 \mid (H\text{-}Ch)_2\right] = 2/3$. From this we obtain, $f(H\text{-}Ch^*)_2 = (1/2)f(H\text{-}Ch')_2$.

For D: The consideration which applies to $(H\text{-}Ch)_2$ applies here. One-third of the time a subject who holds the D system and who receives the feedback sequence $-_1\ +_2\ -_3$ will have the correct dimension second on his list. Therefore

$P(D^* \mid D) = 1/3$, $P(D' \mid D) = 2/3$, and

$f(D^*) = (1/2)f(D')$.

The appropriate corrected equations are

$$f(H\text{-}Ch)_2 = f(H\text{-}Ch')_2 + (1/2)f(H\text{-}Ch')_2$$

$$f(D) = f(D') + (1/2)f(D'), \text{ and}$$

$$f(Fo) = f(Fo') - \left[(1/2)f(H\text{-}Ch^1)_2 + (1/2)f(D')\right] .$$

Appendix I

Percentage of Dimension Checking Problems in Which the Second or Third Hypothesis is Inconsistent with First Feedback

	Third grade	Eighth grade
Four-Dimension-No-Cue	65.9	59.2
Four-Dimension-Cue	52.6	72.9
Eight-Dimension-No-Cue	72.9	63.8
Eight-Dimension-Cue	67.7	63.9
Total	65.3	64.8

Percentage of Dimension Checking Problem in Which the Third Hypothesis is Inconsistent With Second Feedback

Four-Dimension-No-Cue	34.1	45.8
Four-Dimension-Cue	52.6	27.1
Eight-Dimension-No-Cue	44.3	50.0
Eight-Dimension-Cue	33.3	54.1
Total	44.3	45.1

Appendix J

Globally Consistent Problems as a Function of Feedback

Feedback Patterns

Third Grade	++-	-+-	+--	---
Four-Dimension-No-Cue	34 (16/47)	16.7 (8.48)	34.7 (17/49)	22.9 (11/48)
Four-Dimension-Cue	45.8 (22/48)	29.2 (14/48)	35.4 (17/48)	29.2 (14/48)
Eight-Dimension-No-Cue	27.1 (13/48)	25.0 (12/48)	10.4 (5/48)	10.4 (5/48)
Eight-Dimension-Cue	27.1 (13/48)	14.6 (7/48)	8.3 (4/48)	4.2 (2/48)
Eighth Grade				
Four-Dimension-No-Cue	66 (31/47)	56.3 (27/48)	45.8 (22/48)	42.9 (21/49)
Four-Dimension-Cue	64.6 (31/48)	45.8 (22/48)	54.2 (26/48)	29.2 (14/48)
Eight-Dimension-No-Cue	43.5 (20/46)	40.8 (20/49)	46.9 (23/49)	18.8 (9/48)
Eight-Dimension-Cue	45.8 (22/48)	43.8 (21/48)	33.3 (16/48)	23.0 (11/48

Appendix K

Codes for Post-Experimental Verbalization Themes

Code	Theme
E	Elimination
M1	Memory for first trial information
M	Memory
P	Attend to positive instances
PH	Memory for past hypotheses
D	Test by dimensions
F	Use feedback
H	Test hypotheses
DK	Don't know
U	Idiosyncratic
C	Cue card
RR	Retain relevant hypotheses

Categorization of Post Experimental Verbalizations by Theme

Theme	Third grade					Eighth grade				
	4DNC	4DC	8DNC	8DC	Total	4DNC	4DC	8DNC	8DC	Total
E	1	0	0	1	2	14	10	10	11	45
E,M1	0	0	0	0	0	11	7	6	6	30
M	11	9	4	9	33	19	20	17	13	69
M1	4	4	1	3	12	13	13	11	6	43
P	5	6	4	5	20	3	2	1	5	11
PH	2	0	0	2	4	3	1	1	0	5
D	0	0	0	0	0	2	1	2	4	9
F	3	6	3	3	15	0	0	3	0	3
H	5	2	2	1	10	0	0	1	1	2
DK	4	2	3	5	14	0	0	0	0	0
C*	-	2	-	5	7	-	1	-	2	3
C	-	6	-	9	15	-	2	-	5	7
RR*	-	-	9	-	9	-	-	0	-	0
RR	-	-	13	-	13	-	-	7	-	7

Note: *= main theme

	Categories	Theme	Third Grade				Eighth Grade			
			4DNC	4DC	8DNC	8DC	4DNC	4DC	8DNC	8DC
11.	Eliminate Hypotheses	E	0	0	0	0	0	0	1	1
12.	Eliminate	E	0	0	0	0	0	0	1	1
13.	Retain 1st TrialPositive Try Each Hypothesis	M1	1	0	0	0	2	5	2	0
14.	Retain 1st Trial Positive Compare Each Positive	P,M1	0	1	0	0	0	0	0	1
15.	Retain 1st Trial Positive Compare	M1	0	1	0	1	0	0	3	0
16.	Retain 1st Trial Positive	M1	3	2	1	2	2	1	0	0
17.	Compare Positives	P,M	0	1	0	1	0	0	1	0
18.	Remember Positive	P,M	5	2	2	1	1	1	0	0
19.	Look at Positives Remember Past Verbalized Hypotheses	M,P, PH	0	0	0	1	0	0	0	0
20.	Remember Past Verbalized Hypotheses	M,PH	2	0	0	1	2	0	0	0
21.	Remember Past Events. Compare	M	0	0	0	0	0	4	2	4
22.	Remember Past Events	M	0	0	1	0	0	1	1	0
23.	Memory. Vague	M	0	2	0	1	0	0	0	0

Categories	Theme	Third Grade				Eighth Grade			
		4DNC	4DC	8DNC	8DC	4DNC	4DC	8DNC	8DC
24. Dimension Checking	D	0	0	0	0	1	1	1	3
25. Looked at Positive	P	0	2	2	1	1	0	0	2
26. Used Positive Feedback	F	3	5	1	2	0	0	3	0
27. Used Negative Feedback	F	0	1	2	1	0	0	0	0
28. Tried All Hypotheses	H	2	1	0	1	0	0	1	0
29. Unsystematic Hypothesis Testing	H	3	1	2	0	0	0	0	1
30. Don't Know	DK	4	2	3	5	0	0	0	0
31. Idiosyncratic	U	0	1	1	0	1	0	0	0
32. Used Cue Primary	C	-	2	-	5	-	1	-	2
33. Used Cue	C	-	4	-	4	-	1	-	3
34. Remember Relevant	RR	-	-	6 (3/3)	-	-	-	2 (0/2)	-
35. Hard to Remember Relevant	RR	-	-	6 (5/1)	-	-	-	3 (0/3)	-
36. Rehearse Relevant	RR	-	-	1 (1/0)	-	-	-	2 (0/2)	-

Data on Post Experimental Verbalizations

	Categories	Theme	Third Grade				Eighth Grade			
			4DNC	4DC	8DNC	8DC	4DNC	4DC	8DNC	8DC
1.	Retain 1st Trial Positive. Focus	E,M1	0	0	0	0	2	2	1	3
2.	Memorize 8 Hypotheses Preceding First Trial	E,M1	0	0	0	0	1	0	0	0
3.	Retain 1st Trial Positive Information Eliminate Each Feedback Trial	E,M1	0	0	0	0	4	3	1	0
4.	Retain 1st Trial Positive Information Eliminate Wrongs	E,M1	0	0	0	0	1	1	1	1
5.	Retain 1st Trial Positive Information Eliminate Hypotheses not on Positive	E,M1	0	0	0	0	2	0	1	1
6.	Retain 1st Trial Positive Eliminate	E,M1	0	0	0	0	1	1	2	1
7.	Retain Positive Eliminate not on Positive	E,M,P	0	0	0	1	1	1	0	2
8.	Eliminate Hypotheses on Negative	E	1	0	0	0	0	1	0	0
9.	Eliminate Dimensions	E,D	0	0	0	0	1	0	1	1
10.	Eliminate Past Verbalized Hypotheses	E,PH	0	0	0	0	1	1	1	0

Examples
Post Experimental Verbalizations

Number	Group	Verbalization
1.	8th grade 8DNC	I rehearsed the four cues on the first card that could be right. On the second card I got rid of two and then on the third, I knew the answer.
2.	8th grade 4DNC	Process of elimination. Take the two figures-memorize then keep the four ideas on the right one. Then each time eliminate what's wrong.
3.	8th grade 4DNC	Elimination. Picked one-if right-memorized- if one was wrong, three other possibilities-then two possibilities- sometimes after the second, only one.
4.	8th grade 4DNC	Process of elimination. If picked wrong, look at right card . Tried to memorize first card that was right and then eliminate if they were on wrong cards.
5.	8th grade 4DNC	Process of elimination. Looked at first card saw what was right each time. Eliminate the ones that weren't on the right cards.
6.	8th grade 4DC	I remembered all the things that were on the right card and eliminated what was wrong.
7.	3rd grade 8DC	Everytime before you turned a page, I looked at right one and said it over and over again in my mind- either-or- and then stopped saying when I knew it was wrong.
8.	3rd grade 4DNC	First I chose one than I chose another. I think back. I remember which was wrong. I take the things that have been on the wrong one out and then I try the ones that were on the right ones. I tried them all.
9.	8th grade 4DNC	First choose one thing like size -tried big and small and then eliminated that and tried another thing.
10.	8th grade 4DNC	I used the process of elimination- eliminate those I used -but sometimes I forgot.
11.	8th grade 8DC	Looked at all things-picked one that had some of the things- if it was wrong, eliminate those characteristics.
12.	8th grade 8DNC	By the process of elimination. Sometimes I wasn't listening when you told me what it could be and I coudn't remember.

Number	Group	Verbalizations
13.	3rd grade 4DNC	I remembered first four right ones. I tried one than another. I kept those in mind and then fourth answer had to be right.
14.	8th grade 8DC	Tried to remember first right one and looked at cards that were right for the same ones as the first one.
15.	3rd grade 4DC	I picked a letter- if it was right, I picked the same thing. If it was wrong I would see if there was one that was right that was the same as the first right card.
16.	3rd grade 4DC	I tried to remember the first card that was right.
17.	3rd grade 4DC	When it was wrong, I looked at the other one. I picked one from the other that had been right-tried to remember what had been right.
18.	3rd grade 4DNC	I tried to remember what was right on the others and then I tried them. When you said it wasn't, I tried another one.
19.	3rd grade 8DC	I tried to see what all the right ones had. Then everytime the last one. The ones I had already picked, I didn't pick any more.
20.	3rd grade 8DC	I tried to remember what I said.
21.	8th grade 4DC	I tried to figure it out. If wrong, I tried another. I tried to remember the card that was right just before.
22.	8th grade 4DC	I remembered the one before that was right and tried to pick what they had in common.
23.	3rd grade 8DC	Used the cards and narrowed it down. I tried to remember but I forgot.
24.	8th grade 8DNC	I did small, if not that-did large--if not this, than that, one set at a time. Sometimes I forgot what counted.
25.	8th grade 8DC	I picked any one at first. I looked at the right card. If it didn't have it, I didn't use it again. If it did I used it again. If not, I went to another thing on the right picture.
26.	8th grade 8DNC	I took one at random. I'd try it again if it was right.
27.	3rd grade 4DC	I just guessed. When I kept seeing those different pictures, I knew what it was. If it was wrong, I tried another one until I got it.

Number	Group	Verbalization
28.	8th grade 8DNC	I tried each one.
29.	8th grade 8DC	I just picked around- anything I could think of.
30.	3rd grade 8DC	I don't know
31.	8th grade 4DNC	I was thinking of all the possibilities- I tried certain things- I summed them up -like I'd try T and Filled In. If that wasn't right, I had to alternate.
32.	8th grade 8DC	I went through the list. I tried each one.
33.	8th grade 8DC	I used the card. When the one I chose was wrong, I picked another one that was on the right picture.
34.	3rd grade 8DNC	I remembered the ones you said.
35.	3rd grade 8DNC	Sometimes I forgot what you told me in the beginning.
36.	3rd grade 8DNC	I said them over and over again and tried to remember which could be right.

References

Appel, L.F., Cooper, R.G., McCarrell, N., Sims-Knight, J. Yussen, S.R., & Flavell, J. H. The development of the distinction between perceiving and memorizing. Child Development, 1972, 43, 1365-1381.

Belmont, J.M., & Butterfield, E.C. What the development of short term memory is. Human Development, 1971, 14, 236-248.

Broadbent, D. Perception and communication. New York: Pergamon Press, 1958.

Bruner, J.S., Goodnow,J.J.,& Austin, G.A. A Study of thinking. New York: John Wiley & Sons, Inc. 1956.

Collins, J.W., Learning of media content: a developmental study. Child Development, 1970, 41, 1133-1142.

Corsini, D.A. Memory: interaction of stimulus and organismic factors. Human Development, 1971, 14, 227-235.

Doyle, A. Listening to distraction: a developmental study of selective attention. Journal of Experimental Child Psychology, 1973, 15, 100-115.

Druker,J.F. & Hagan, J.W. Developmental trends in the processing of task-relevant and task-irrelevant information. Child Development, 1969, 40, 371-382.

Egeth, H. Selective attention, Psychological Bulletin, 1966, 67, 41-57.

Eimas, P.D. A developmental study of hypothesis behavior and focusing. Journal of Experimental Child Psychology, 1969, 8, 160-172.

Eimas, P.D. Effects of memory aids on hypothesis behavior and focusing in young children and adults. Journal of Experimental Child Psychology. 1970, 10, 319-336.

Falmagne, R. Construction of a hypothesis model for concept identification. Journal of Mathematical Psychology. 1970, 7, 60-96.

Flavell, J.H. The developmental psychology of Jean Piaget. New York: D. Van Nostrand Co. Inc. 1963.

Flavell, J.H. First discussant's comments: what is memory development the development of. Human Development, 1971, 14, 272-277.

Freeman, M.F. & Tukey, J.W. Transformations related to the angular and the square foot. Annals of Mathematical Statistics, 1950, 21, 607-611.

Gholson, B. & Danziger, S. Effects of two levels of stimulus complexity upon hypothesis sampling systems among second and sixth grade children. Journal of Experimental Child Psychology, 1975, in press.

Gholson, B., Levine M., & Phillips, S. Hypotheses, strategies and stereotypes in discrimination learning. Journal of Experimental Child Psychology, 1972, 13, 423-466.

Gholson, B. & McConville, K. Effects of stimulus differentiation training upon hypotheses, strategies, and stereotypes in discrimination learning among kindergarten children. Journal of Experimental Child Psychology. 1974, 18, 81-97.

Gholson, B. & O'Connor, J. Dimensional control of hypothesis sampling during three-choice discrimination learning. Child Development, 1975, 46, 894-903.

Gholson, B., Phillips, S. & Levine, M. Effects of the temporal relationship of feedback and stimulus information upon discrimination learning strategies. Journal of Experimental Child Psychology, 1973, 15, 425-441.

Gumer, E. & Levine, M. M. The missing dimension in concept learning: dimensionality or local consistency. Journal of Experimental Psychology, 1971, 90, 39-44.

Hagan, J. W. The effect of distraction on selective attention. Child Development, 1967, 38, 685-694.

Hagan, J.W. Some thoughts on how children learn to remember. Human Development, 1971, 14, 262-271.

Hagan, J.W. Strategies for remembering. In S. Farnham-Diggory (Ed.) Information processing in children. New York: Academic Press, 1972, 66-79.

Hagan, J.W., & Frisch, S.R. The effect of incidental cues on selective attention. Unpublished manuscript. University of Michigan, 1968.

Hagan, J.W. & Hale, G.H. The development of attention in children. In A.D. Pick (Ed.) Minnesota Symposium Child Psychology. Vol. 7. Minneapolis: University of Minnesota Press, 1973, pp 117-140.

Hagan, J.W., Meacham, J.A., & Mesibov, G. Verbal labeling, rehearsal and short term memory, Cognitive Psychology. 1970, 1, 47-58.

Hagan, J.W. & Sabo, R.A. A development study of selective attention. Merrill-Palmer Quarterly, 1967, 13, 159-172.

Hale, G.A., Miller, L.K. & Stevenson, H.W. Incidental learning of film content: a developmental study. Child Development, 1968, 39, 69-78.

Ingalls, R.P., & Dickerson, D.J. Development of hypothesis behavior in human concept identification. Developmental Psychology, 1969, 1, 707-716.

Karpf, D., & Levine, M. Blank trial probes and introtacts in human discrimination learning. Journal of Experimental Psychology, 1971, 90, 51-55.

Kemler, D.G. A developmental study of hypothesis-testing in discrimination learning tasks. Unpublished doctoral dissertation. Brown University, 1972.

Kendler, H.H. & Kendler, T.S. Vertical and horizontal processes in problem solving. Psychological Review, 1962, 69, 1-16.

Levine, M. Hypothesis behavior by humans during discrimination learning. Journal of Experimental Psychology, 1966, 71, 331-338.

Levine, M. A cognitive theory of learning. New Jersey: Lawrence Erlbaum Associates, 1975.

Maccoby, E.E., & Hagan, J.W. Effects of distraction upon central versus incidental recall: developmental trends. Journal of Experimental Child Psychology, 1965, 2, 280-289.

Moely, B.M., Olson, F.S., Halwes, T.G. & Flavell, J.H. Production deficiency in young children's clustered recall. Developmental Psychology, 1969, 1, 26-34.

Moray, N. Attention: selective processes in vision and hearing. New York: Academic Press, 1970.

Mood, A.M. Introduction to the theory of statistics. New York: McGraw, Hill, 1950.

Neimark, E.D. & Lewis, N. The development of logical problem-solving strategies. Child Development, 1967, 38, 107-117.

Neimark, E.D. & Lewis, N. Development of logical problem solving: A one year retest, Child Development, 1968, 39, 527-536.

Neimark, E., Slotnick, N.S. & Ulrich, T. Development of memorization strategies. Developmental Psychology, 1971, 5, 427-432.

Neisser, U. Cognitive psychology. New York: Appleton, Century, Crofts, 1967.

Neussle, W. Reflectivity as an influence on focusing behavior of children. Journal of Experimental Child Psychology, 1972, 14, 265-276.

Petronovich, L.F. & Hardyk, C.D. Error rates for multiple comparison methods. Some evidence concerning the frequency of erroneous conclusions. Psychological Bulletin, 1969, 71, 43-54.

Phillips S., & Levine, M. Probing for hypotheses with adults and children: Blank trials and introtacts. Journal of Experimental Psychology: General, 1975, 104, 327-354.

Piaget, J., & Inhelder, B. The psychology of the child. New York: Basic Books, Inc. 1969.

Sabo, R.A. & Hagan, J.W. Color cues and rehearsal in short-term memory. Child Development, 1973, 44, 77-82.

Siegal, A., & Stevenson, H.W. Incidental learning: a developmental study. Child Development, 1966, 37, 811-817.

Simon, H.A. On the development of the processor. In S. Farnham-Diggory (Ed.) Information processing in children. New York: Academic Press, 1972.

Tighe, L.S. & Tighe, T.J. Transfer from perceptual pretraining as a function of number of task dimensions. Journal of Experimental Child Psychology, 1969, 8, 494-502.